CESifo WORKING PAPERS

10268
2023

February 2023

Firm Training

Dan A. Black, Lars Skipper, Jeffrey A. Smith

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10268*

Firm Training

Abstract

Workers acquire skills through formal schooling, through training provided by governments, and through training provided by firms. This chapter reviews, synthesizes, and augments the literature on the last of these, which has languished in recent years despite the sizable contribution of firm training to the overall stock of worker human capital. We engage with research on the determinants of receipt of firm training, the effects of firm training on workers outcomes, and various policy debates related to firm training, including training taxes, training subsidies, non-compete agreements, and the minimum wage. Our discussion emphasizes the complex measurement issues associated with firm training and the interplay of applied theory and applied econometrics in the related empirical literature.

JEL-Codes: I200, I240, J240, J420, J310.

Keywords: training, human capital, firm, worker, classroom, learning by doing, monopsony, minimum wage, training tax, non-compete.

Dan A. Black
Harris School of Public Policy
University of Chicago
USA – Chicago, IL 60637
danblack@uchicago.edu

Lars Skipper
Department of Economics
Aarhus University
Denmark - 8210 Aarhus V
lskipper@econ.au.dk

Jeffrey A. Smith
Department of Economics
University of Wisconsin-Madison
USA - Madison, WI 53706-1393
econjeff@ssc.wisc.edu

January 25, 2023
We thank the editors for their enthusiasm and patience, Jacob Klerman for discussions of the sectoral training literature, Lance Lochner for help with the literature on credit constraints in higher education, Amalie Borre Berg for helpful research assistance on PIAAC, Sarah Stith for pointers to the learning-by-doing literature, Evan Starr for guidance on the literature on non-compete agreements, and Lois Miller for a careful reading.

1 Introduction

1.1 Why study firm training?

In his presidential lecture to the American Economic Association entitled "Investment in Human Capital," Theodore Schultz (1961) states:

> I shall concentrate on five major categories: (1) health facilities and services, broadly conceived to include all expenditures that affect the life expectancy, strength and stamina, and the vigor and vitality of a people; (2) on-the-job training, including old-style apprenticeship organized by firms; (3) formally organized education at the elementary, secondary, and higher levels; (4) study programs for adults that are not organized by firms, including extension programs notably in agriculture; (5) migration of individuals and families to adjust to changing job opportunities.

In regard to item (1), a vast literature inside (and outside) health economics studies the effects of investments in health capital via e.g. diet, exercise, and education. Grossman(2000) summarizes some of this literature. In regard to item (3), a vast literature studies all aspects of elementary, secondary and tertiary education, including most of the chapters in this volume and in earlier volumes of the *Handbook of the Economics of Education*. In regard to item (4), a vast literature, summarized in McCall et al. (2016), considers the organization and effects of "study programs not organized by firms", mainly in the context of active labor market programs provided by governments to unemployed (or sometimes under-employed) workers. Finally, in regard to item (5), a vast literature in demography and in labor economics studies the determinants and "returns" to migration; see the summaries in Borjas (1999), Chiswick and Miller (2015), and Jia et al. (forthcoming).

This chapter reviews (and augments) the not-at-all-vast literature related to Schultz's item (2), on-the-job training. In particular, we engage with the literature on training provided by employers to workers in the developed world. Though firms may receive government subsidies for the training they provide, the training we consider differs from that provided by active labor market programs in that the recipients already have a job with a firm, and the training aims to improve their performance with that firm. The training we consider ranges from highly structured classroom instruction over a period of days, weeks, or months that ends with a generally recognized credential, to one worker answering another worker's questions about the local management information system while leaning over the divider that separates their desks. We largely exclude the apprenticeships that Schultz took the time to mention as Wolter and Ryan (2011) and, more recently, Carruthers and Jepsen (2021) ably survey that literature. For reasons explained in Section 2, we also exclude the modest literature on learning-by-doing, which lies at a surprising theoretical and empirical distance from the bulk of the firm training literature. Finally, we also exclude the small literature on firm training in the developing world, see e.g. Rosholm et al. (2007), both to keep our scope manageable and because of the important institutional differences between firms in the developing and developed worlds.

The relative lack of academic studies on firm training has several related causes, causes that our chapter makes plain. That set of causes decidedly does not include either a lack of substantive importance or a lack of policy importance. One path to understanding this starts with the estimates of direct training costs and the opportunity costs of worker time spent in training that we present in Section 4. The inputs of time and money devoted to training on the job by firms and workers imply substantial economic value-added, perhaps on the same scale as that created by formal schooling. Put differently, the volume of investment in firm training suffices to make it a major component of the overall stock of human capital.

Though too small, in our view, relative to the substantive economic importance of the topic, the economic literature on firm training has made substantial headway both theoretically and empirically since its initiation about 65 years ago. Our chapter documents that progress. It also highlights where and how researchers can contribute to the firm training literature at the margin. In a couple of cases, we could not resist the temptation to engage with open research questions ourselves. The empirical results from those engagements appear in Sections 4.4 and 5.4.

Several themes recur throughout the chapter. The first concerns the importance of definition and measurement issues. We group the two because one important source of measurement difficulties is the sublime level of empirical heterogeneity in firm training. Defining and grouping training types thus represents a more pressing concern than in adjacent domains such as primary and secondary (and to some extent tertiary) education. In most countries, primary and secondary schools share key institutional features (e.g. the number of years to completion of each stage) and lead to a small number of familiar credentials. This commonality and familiarity eases the problem of measurement on surveys both in the sense of making it simpler to ask the questions and in the sense of making it simpler for respondents to answer them.

A second recurring theme in the firm training literature centers on the interplay of theory and measurement. Relative to many empirical literatures within labor economics and the economics of education, theory plays a key role in the conversation around firm training, and does so despite the predominance of "design-based" rather than "structural" empirical work. A final theme concerns the importance of resurrecting the literature on firm training within economics. Researchers sometimes fall into a trap of assuming that things they have trouble getting a handle on must therefore lack substantive importance. General equilibrium effects, as in Lise et al. (2004), and the value of program participants' time, as in Greenberg and Robins (2008), provide two examples of this trap from the literature on active labor market programs. We worry that something similar, a bit of "out of sight, out of mind", applies to firm training within labor economics. Our chapter pushes back against that bad mental habit by recollecting the economic importance of firm training and by pointing out opportunities to move forward.

1.2 What we do

We start our journey into the economic literature on firm training with typologies of training in Section 2. Definitional issues haunt this literature, and the reader should pay close attention to slippery terms like "formal" and "informal" throughout. Difficulties of measurement bear some of the blame for the lethargic pace of research on training over the past three decades. Section 3 lays out the measurement issues conceptually and reviews the small empirical literature that attempts to quantify the nature and extent of measurement error in survey data on firm training. In Section 4, we review studies of the nature, extent, and determinants of firm training. These studies describe the characteristics of firm training participants and document, in some cases, the conditional probability of training of particular types and duration as a function of worker, establishment, and/or firm characteristics. We also undertake our own study of firm training incidence using the data from the recent Programme for the International Assessment of Adult Competencies (PIAAC) survey, which produces comparable data on (among other things) receipt of firm training in a variety of developed countries.

Section 5 follows the theoretical literature on training over time starting with the seminal contributions of Becker and Mincer in the early 1960s and continuing through more recent developments related to imperfectly competitive labor markets and multi-dimensional skills. We intersperse the related empirical work with the theory because the literature does so as well. Indeed, much of this literature seeks to learn about training indirectly from the traces it leaves behind in starting wages, wage-tenure profiles, and other labor market outcomes. The concluding part of Section 5 investigates the small literature that directly estimates the treatment effect of training. Section 6 confronts the policy literature on training. It starts with a review of the rationales offered in the literature for government intervention in this market. It proceeds with discussions of training taxes, subsidies, and vouchers, and of the research examining the intersection of firm training with minimum wage policy, with policy regarding non-compete agreements, and with active labor market policy via sectoral training programs. Section 7 revisits our main points and offers a final call to revive the literature on firm training.

2 Typologies of training

As noted in the introduction, definitions matter in the literature on firm training. Consider two broad definitional issues to start our discussion. The first concerns the extensive margin definition of training; this definition draws the boundary between activities that constitute training and activities that do not. As we read the literature, researchers generally agree on this aspect of the problem of defining training. The second concerns the intensive margin definition of types or categories of training. Here the literature evinces an enthusiastic heterogeneity; if not a training type zoo to match the "identification zoo" of Lewbel (2019), we at least possess a training type menagerie.

The would-be constructor of a training typology confronts the multi-dimensional nature of firm training. Should the types build on the duration of training? Should they build on the purpose of training, say team-building or learning new skills or checking regulatory boxes? Should the types build on whether the firm provides the training directly or hires an outside provider? We will show in Section 3 that each of these dimensions receive some attention in at least one of the surveys we consider. Having chosen a dimension or two on which to base a typology, the question then arises of its optimal coarseness. Should it distinguish between classroom training that occurs on-site at the firm or off-site with a provider, or just pool the two in a single category? Should it distinguish between casual training from supervisors and that from co-workers? And so on.

As our primary typology we adopt the one defined by the firm training questions in the PIAAC, the ongoing survey operated by the OECD that seeks to measure education and skills in comparable ways across countries. We will have much more to say about the PIAAC later in the chapter. For now, what matters is that the PIAAC typology partitions firm training into four types or categories: formal, informal, casual, and learning-by-doing. As we explain, it relies not on a single dimension of the training experience to define its types but rather on multiple dimensions.

The first PIAAC training type, "formal training" includes training ("study" in the words of the PIAAC survey question) aimed at a formal qualification, whether part-time or full-time. Formal credentials include both academic credentials such as associate's degrees, or more modest, vocationally-oriented credentials such as the many certificates offered by Microsoft that document proficiency in particular skills or software packages.

The second PIAAC training type, "informal training" captures a wide variety of activities, including "courses conducted through open or distance education", "organized sessions for on-the job training", or "seminars or workshops", where the quoted text comes from the U.S. version of the PIAAC survey instrument. The presence or absence of a formal qualification at the end distinguishes formal from informal training in the PIAAC scheme.

The third PIAAC training type, "casual training" builds on a survey question that asks: "In your own job, how often do you learn new work-related things from co-workers or supervisors?". This question comes after the informal and formal training questions, apparently with the intention that a formal course taught by a supervisor or co-worker would count as informal training rather than casual training. Rough distinctions between casual training and formal or informal training include the one-on-one interaction and spontaneity implicit in the casual training question, and the narrow focus on a supervisor or co-worker as the one transmitting the skill.

The fourth and final PIAAC training type comprises "learning by doing" wherein repetition itself leads to improved performance in a work task. Put differently, learning-by-doing differs from the other categories of training because it does not imply forgone work hours. Instead, the training inheres in the work itself. Learning-by-doing has its own specialized literatures, quite distinct from the literatures we survey in the sections to follow. Given the general lack of overlap, we place learning-by-doing outside the scope of our chapter (other than a brief appearance in Section 4.4) and confine ourselves to providing some pointers to the literature.

One learning-by-doing literature focuses on learning at the level of the team or individual, with the individual learning mostly implicit. These studies tend to focus on specific contexts that feature clear quantitative measures of output and/or output quality and accurate measures of the amount of task repetition. Stith (2018), which studies learning-by-doing in transplant surgery, offers a compelling example of this genre. Another literature devotes itself to sorting between learning-by-doing and other kinds of training in the context of life-cycle earnings profiles. Heckman et al. (2002) provide a link into this literature.

We adopt the PIAAC definitions for several reasons. We think four categories represents a reasonable balance in the trade-off between the nuance offered by additional categories and the simplicity of a relatively coarse categorization. The PIAAC typology closely resembles that in other major surveys in this domain such as the European Adult Education Survey (AES) and the Canadian Adult Education and Training Survey (AETS). We say more about these (and other) surveys in Section 3. Finally, given that we make use of the PIAAC data, and thus the PIAAC typology, in our original empirical analysis of the incidence, duration, and determinants of training around the OECD in Section 4.4, it seems natural to adopt it as our baseline.

We have two main frustrations with the PIAAC training typology. The first arises from the fact that exactly the same course could count as either formal or informal training depending on the circumstances. Consider a two-course sequence that leads to some worthy credential. For a worker who completes the first course but not the second, the first course represents informal training because it does not lead to a credential. For a different worker who does complete both courses and obtains the credential, exactly the same course counts as formal training. The second frustration concerns the partial disconnect between everyday notions of formality and the informal label applied to courses or seminars not leading to a credential. An informal course as PIAAC defines it could nonetheless seem quite formal to its participants, depending on its level of organization and mode of presenting the material. Despite these frustrations, the PIAAC struck us as the best of the available typologies, and we wanted to avoid cluttering the literature with yet another typology of our own, particularly given that it would not correspond to any of the surveys whose data appear in our chapter.

We refer to these definitional issues throughout the chapter. In Section 3 we link them to the dilemmas of training measurement. In Section 4 they will help to account for the sometimes remarkably wide differences in empirical findings related to firm training across studies. In Section 5, they will interact with the applied theory of firm training in interesting and important ways while in Section 6 they will pop up again in our discussion of training taxes. We end this section with a repeat of our recommendation that readers of the firm training literature, both those encountering it for the first time and those who have engaged with it before, must keep these definitional questions mentally front and center in a way not required by most other applied economics literatures.

3 Measuring firm training

3.1 Introduction

This section describes what we know about the measurement of training in surveys. This matters because nearly all of our empirical knowledge regarding the incidence, nature, and extent of firm training derives from responses to surveys administered to workers, to their managers, or to human resources staff at firms. Perhaps not surprisingly, firms do not routinely share their administrative data on training (which would likely cover only more formal types in any event) and governments do not maintain any sort of administrative register data on firm training.

Section 3.2 describes the extant variation in survey measures of training, noting key dimensions of variation and foreshadowing the resulting variation in estimates of training incidence and duration that we dissect in Section 4. Section 3.3 considers the very limited evidence on measurement error in survey data on training, which comprises a handful of papers that compare training reported in worker surveys with training for the same workers reported in employer surveys or in administrative data. Section 3.4 covers the even-more-limited indirect evidence on measurement error, which comes in one case from instrumental variables and in another from comparisons of survey and administrative data on training, not at firms but in active labor market programs. A refrain throughout the section concerns the distinction between measurement error, defined as the difference between what a survey question asks for and what the respondent reports, and measurement differences, defined as the difference between a correct response to one survey question wording and a correct response to an alternative survey question wording. Both play important roles in generating the variation we observe.

3.2 Survey questions about firm training

[INSERT TABLE 1 HERE]

Table 1, drawn from Kristensen et al. (2022), presents questions about firm training drawn from major surveys in the developed world. While surely not the universe of survey questions about training (even in those parts of the world), based on our reading it successfully captures the variation in question style and content. We organize the table by country or country group. Each entry provides the survey name, the wording of the survey measure or measures in English, and sometimes some additional notes in parentheses. Table 1 documents substantial variation in the ways that researchers ask workers and employers about training. The variation in estimates of training incidence, duration, and determinants that results (in part) from this variation in survey measures represents one of the major themes of this chapter. Table 2, also drawn from Kristensen et al. (2022), documents other features of these survey measures. We describe and engage with it more deeply in Section 4.

The survey measures of training described in Table 1 differ on several dimensions. We group these dimensions into three categories: first, details of the survey question or questions; second, the broader survey context that embeds the questions about training; third, the target population for the survey.

Within the category of details of the specific survey questions, we observe three quite distinct genres of survey questions: (1) direct questions about individual training spells or episodes as in the NLSY-79; (2) questions that ask the respondent to accumulate all training episodes over a specific time period, as with the questions from the Equal Opportunity Pilot Program (EOPP); and (3) questions that ask about the amount of training required to reach some notion of proficiency, e.g. "fully trained and qualified" in a job as in the Panel Study of Income Dynamics (PSID) in 1976. Asking the respondent to implicitly combine multiple episodes increases cognitive burden, which may lead to errors. On the other hand, respondents with many episodes may respond to a survey that asks multiple questions about each episode by omitting some episodes in order to reduce the time and effort cost of survey completion. We worry that the proficiency-based formulations add noise to the measure due to heterogeneous interpretations of this admittedly subjective criterion.

A second detail concerns the reference period over which training is measured. Examples from our surveys include the first three months of employment, since the worker was hired, the last three months, the last 12 months, and the last 10 years. Shorter reference periods should make for more accurate responses but will also lead to more respondents reporting no training, given its episodic nature. Asking about the (likely quite salient) time since hiring rather than using a reference period based on calendar time may reduce respondent burden and thus measurement error.[1]

A third detail concerns respondent interpretation of the meaning of the general term "training" or of more specific terms such as "informal training" or "credential" in light of the cues they receive from the survey instrument and/or an interviewer. The question wording and other cues will affect the respondent's *comprehension* of the question, i.e., their process of parsing the question, identifying and relating the components of the question to each other, assigning meaning to its key elements ("stu-

[1]The European Community Household Panel (ECHP) created its own recall period problem by asking about training since the preceding January and then failing to consistently field the survey early in the next calendar year as originally planned. In practice, survey timing varied widely across countries, implying substantial variation in the realized recall period as well as the possibility of overlapping recall periods in adjacent survey years. See Peracchi (2002) for the sordid details.

dents", "apprentice", and "training"), inferring the purpose behind the question, and determining boundaries and any overlap among permissible answers. For example, the respondent may ponder (consciously or not consciously) what separates a formal training session from an informal one or what distinguishes a course and a workshop. Looking through the different surveys in our Table 1 and the questions on training therein, we can easily imagine how respondents might have trouble making some of the distinctions they embody, as most respondents will not have spent time reading the training literature.[2]

Turning now to the broader survey context, we first remark that the questions in Table 1 differ in the path that the respondent follows to get to them, i.e. the routing within the survey that causes a particular respondent to reach or not reach the specific training questions embedded in the survey. For example, the AES asks respondents about informal training episodes first and then asks about the reasons for the informal training episodes. In contrast, the NLSY-79 1993 and 1994 surveys (rounds 15 and 16 in their terminology) first ask whether or not the respondent needed new job skills in the past 12 months (due to e.g. new products or services, new equipment, reorganizations, new government regulations, etc.). Only a positive response to this gateway question sends the respondent to the questions about informal training; a negative response implies a valid skip. We expect the NLSY routing to yield a lower incidence rate, all else equal.[3]

Another contextual difference centers on whether the training questions form part of a survey instrument narrowly focused on education and training (such as AES, AETS, the Continuing Vocational Training Survey (CVTS), and the Further Training as a Part of Lifelong Learning (WeLL)) or form part of a broader, general survey (such as the

[2] Belson (1986) suggests that respondents may assign a range of meanings even to seemingly straightforward terms like the "you" in the questions on training participation in Table 1; i.e., does the question mean you personally, or you and your spouse, or you and your family, etc.

[3] The NLSY-79 reintroduced questions on informal training for its 1998 to 2000 surveys. The revised instrument dropped the gateway question and changed the time period covered by the questions to the last four weeks rather than the last 12 months. Unlike some parts of the NLSY-79 survey instrument that have remained relatively stable after an initial period of experimentation, the training component has experienced ongoing tinkering. See the NLSY-79 web page on training for more details: https://www.nlsinfo.org/content/cohorts/nlsy79/topical-guide/education/training

German Socioeconomic Panel (SOEP), NLSY, PIAAC, and PSID). This context will matter if, for example, non-respondents in general surveys have a different propensity to participate in training than non-respondents to specialized surveys. From the survey methodology literature we know that respondents' interest in the topic, called *topic saliency*, affects their willingness to respond; see, e.g., Heberling and Baumgartner (1978) and Sheehan and McMillan (1999) for evidence on survey response and Adua and Sharp (2010) for item response. People willing to sit for half an hour and talk about training and education may be more likely to participate in said activities. Item non-response behavior might vary between specialized and general surveys as well. For instance, one can imagine that respondents might more readily pass on questions that do not appear central to a particular survey. Or the interviewer might try harder to obtain responses on questions more central to the overall mission of the survey. Indeed, one might conjecture that, *ceteris paribus*, specialized surveys would realize lower survey response rates but also lower item non-response rates conditional on survey response.

Our third and final category relates to the survey population. The fourth column of Table 2 documents a wide range of populations, from newly hired workers in low paying firms in the EOPP and Small Business Administration (SBA) surveys, to everyone in particular birth cohorts in the NLSY surveys, to workers at establishments above a certain size in the Survey of Employer-Provided Training (SEPT), to all "prime age" adults in the International Adult Literacy Survey (IALS).[4] Because, as we show in Section 4, the incidence of training varies with worker, establishment, and firm characteristics, who gets included has a natural effect on estimated incidence rates. The same point holds with training type and training duration. A somewhat more subtle pathway from sample composition, operating through the definition of a survey's population of interest, to realized training measures runs through the interaction of respondent characteristics and the survey response, item response, and measurement

[4] An establishment is a single business location. Many firms have only a single establishment but others combine multiple establishments. Conceptual complications arise in franchise contexts, such as McDonald's.

error processes. The broader survey literature, e.g. Bound et al. (2001) and Meyer et al. (2015), shows all of these processes to vary with respondent characteristics such as years of schooling, which in turn correlate with training.

We looked for, but did not find, research on the wording of questions about firm training informed by the broader literature on survey measurement. We think such work would add great value. We conclude this subsection with a couple of specific suggestions for research we would like to see inspired by the existing survey measures of firm training captured by Tables 1 and 2.

First, while recognizing that questions about firm training represent but a very modest component of most surveys of which they form a part, going forward we would like to see tests of survey questions that pay more attention to the nature or point of the training in addition to the form of its delivery, its duration, who (nominally) paid for it, and whether or not it leads to some sort of formal credential. We have in mind letting the respondent differentiate between training that aims to teach the individual a new skill (e.g. how to operate a new machine tool or the new commands in Stata), training that certifies the retention of an existing skill (e.g. annual security training for users of restricted data), and training that aims to make a work group more productive without necessarily directly increasing the productivity of any individual worker (e.g. teamwork or diversity training). These different aims of training have very different implications for thinking about advancing on the applied theory front and for studies that seek to estimate impacts of firm training.

Second, we would like to see further research inspired by the training logs kept by workers and managers in the SEPT-95 survey analyzed in detail in Frazis et al. (2000), which we describe in greater detail in Section 4.2.3. We have in mind keeping the spirit but updating the technology, perhaps by replacing the logs with workplace video cameras, as the Gates Foundation's Measuring Effective Teachers (MET) study and the Trends in International Mathematics and Science Study (TIMSS) 1999 Videotape Classroom Study do with teachers in classrooms.[5] Workplace videos would allow re-

[5]The MET videos reside at ICPSR: https://www.icpsr.umich.edu/web/ICPSR/studies/34771, while the

searchers to observe training activities with minimal measurement error and to study alternative coding schemes. Moreover, administering standard survey questions about training like those in Table 1 to the workers and managers in the videos would provide valuable knowledge on the sorts of measurement error they embody.

3.3 Direct validation evidence on survey measures

This section ponders the (very) limited empirical evidence on the quality of survey responses regarding firm training. In this context, quality usually refers to the correspondence between a worker report (always a survey) and a firm report (sometimes a survey and sometimes administrative data). The training measures considered include most of the types in Table 1, such as binary measures of incidence and continuous measures of training hours over particular intervals as well as continuous measures of how long it takes a worker to become fully trained in a particular job. We found and review just three papers in total: Barron et al. (1997a), Krueger and Rouse (1998), and Kristensen et al. (2022). In any validation exercise, a prosaic version of the sublime philosophical question "What is truth?" naturally arises. We note how each study's authors answer this question.

In the background of these validation studies stands a broader validation literature that looks at other features of jobs. The PSID validation study, which administered the PSID survey instrument to workers in a single large firm and compared the responses to the firm's administrative records, has direct relevance, though it did not ask about training. Bound et al. (2001) cover the PSID validation study in their survey on measurement error in surveys. A more distant but also more extensive literature compares alternative measures of earnings from surveys (e.g. measures that ask the respondent directly about total earnings versus measures built up from underlying questions about employment spells, wages, and usual hours) and from various administrative data sources (e.g. in the United States, UI wage records, income tax returns, and social security data). Hotz and Scholz (2002) and Wallace and Haveman (2007)

TIMSS videos reside at https://www.timssvideo.com/

provide fine examples of this genre.[6]

3.3.1 Barron, Berger and Black

Barron et al. (hereinafter BBB) (1997a,b) provide the first (of just two) published comparisons we consider in some detail and the first (of just two) studies that focus primarily on training measurement. They undertake a modest survey of firms and their recently hired workers with the goal of comparing worker and firm responses to exactly the same questions about the on-the-job training the firm provides and the workers receive. They do not view either survey response as necessarily closer to the "truth" of training receipt.

In order to focus their limited survey resources on establishments likely to have a recent hire, they sample from the population of establishments with at least 100 workers. Their survey had a very low response rate of around 20 per cent, when defining a response as both a firm and its worker completing at least one interview (out of three the firms could do and two the workers could do). In a world with many validation studies of survey measures of firm training, we might discount this study based on its response rate, even though we do not have a strong *a priori* story regarding the association between survey non-response and the similarity of firm and worker responses. But we do not live in such a world.

The interviews were performed in 1993 and concerned themselves with training received during the first two and first four weeks of employment. In particular, BBB (1997a,b) consciously adopted questions similar to those of the EOPP survey shown in Table 1, including the number of hours spent on various types of training and the length of time a new employee in the same job as the survey worker would take to become fully trained and qualified. The survey's hours measure differs in consisting of separate components for weeks and hours per week of each training type, which yields a constructed measure of total hours, whereas the EOPP and SBA surveys ask directly

[6] The modest literature on measurement error in formal schooling also has relevance; see e.g. Black et al. (2003) or Battistin et al. (2011).

about total hours. In addition, the BBS survey asks a variety of questions about other job attributes such as starting wage, benefits, and collective bargaining coverage.

Turning to the findings, the correlations between the worker and establishment responses regarding hours of particular types of training (on-site formal, off-site formal, informal by manager, informal by co-worker, informal by watching) in the first month of employment range from a low of 0.17 to a high of 0.46. Aggregating up into two categories yields correlations of 0.42 for formal training and 0.40 for informal. They, and we, would have expected a larger advantage for formal training as more salient and in some sense better defined. Aggregating to total hours of any kind of training raises the correlation to 0.47. The responses to the question about the number of weeks of training required to become fully trained exhibit a correlation of 0.17. While one expects some slippage on continuous measures such as hours, the low correlations on the incidence of formal training in the first month of employment, equal to 0.32 for on-site training and 0.38 for off-site training, suggest a deeper issue at work in one or both responses. These findings likely signal substantial bias in studies that seek to estimate the causal effect of firm training using survey measures of training, whether provided by the worker or by the firm.

The authors put the remarkable inconsistency they find for the training measures in context by comparing worker and firm agreement on other job aspects in their data. The levels of agreement on these other variables, such as a correlation of 0.97 for the starting wage responses and correlations of 0.77 for hours worked and 0.71 for collective bargaining coverage, make clear that the poor performance on the training variables does not simply result from low response effort by the workers and/or the firms in the BBB (1997a,b) data. Moreover, the levels of agreement found for these other variables resemble those found for the same variables in the PSID validation study, again pointing to issues specific to the training measures as the cause of the poor performance.

In brief, BBB (1997a,b) bring a bevy of bleak findings. While our first impulse loads the measurement error onto the worker reports, some fraction likely belongs to the establishments as well, who may not put much effort into their no-stakes response.

Put differently, in BBB (1997a,b) neither set of responses likely even approximates truth. In terms of external validity, we do not worry too much about the limitation to establishments with at least at least 100 workers nor about the low survey response rates. We do worry about temporal external validity to the present day, as workers, jobs, and firm training technology have all changed substantially over the past three decades.

3.3.2 Krueger and Rouse

Krueger and Rouse (1998) study the effects of a novel, voluntary workplace literacy program offered to low-skilled, hourly workers at two mid-sized companies in New Jersey, one in a service industry and one in a manufacturing industry. The program ran from October 1992 to February 1994. A local community college designed and administered the program, which received support from the firms and from a federal grant. The program featured classes at the work site during paid working hours. Overall, 480 workers attended a total of 12,500 hours of classroom training.

The authors have two measures of participation for about half of their sample: one from attendance records kept by the community college ("truth") and one from a survey they administered to the workers. Among the workers with both measures, 17% of the self-reported trainees did not show up in the administrative records while 6% of the trainees in the administrative records claim not to have received any training in the survey. Given the substantive focus of the paper, the authors do not belabor these findings, but they do note that, considering the administrative data as truth, the use of the self-report data would imply attenuation in impact estimates of up to 28 percent.

3.3.3 Kristensen, Skipper and Smith

In unpublished work, Kristensen et al. (2022) (hereinafter KSS) use Danish administrative data linked to survey data from the AES to compare firm and worker measures of training. They benefit from the two features of the Danish environment. First, the Danish government heavily subsidizes all types of formal education and training for workers.[7] In order to implement the subsidy, firms and training providers work with the government to provide accurate third-party administrative measures of the nature and extent of training received by workers. Second, the Danish research data environment allows KSS to link the administrative data on firm training to Danish respondents from the 2011 and 2016 waves of the AES. Danish response rates to these waves of the AES exceed 50 percent.[8]

Their study has several advantages relative to the other studies considered here. They have all the firms in Denmark, not just one firm, as in the PSID validation study, or two firms, as in Krueger and Rouse (1998), or 305 establishments, as in BBB (1997a,b). The firm measure of training represents an accounting measure rather than a survey response from someone at the firm. Because of the financial stakes that implicit in the subsidy, both the firm and the government have an incentive to avoid errors in the administrative measure. Indeed, KSS treat the administrative data as "truth" in their analysis for this reason. On the survey side, the AES questions that KSS rely on closely resemble those of many other surveys around the world whose data have played major roles in the extant literature on firm training. Also, while one might expect Danish firms to provide more training than firms in other countries with smaller subsidies, they see no particular reason to think this would affect the reporting errors at the center of their investigation.

Of course, subsidizing formal and informal (as defined in Section 2) training but not casual training may lead firms to convert casual training to informal and formal training at the margin in pursuit of the subsidy, a topic that arises again in the context of our discussion of training taxes in Section 6.2. From the measurement angle, moving training from informal to formal affects our comparisons only if formal training gets reported more accurately than informal training, which BBB's (1997a,b) findings suggest is not the case.

[7] Of course, it also subsidizes formal schooling for those preparing for the labor market and spends vast sums on active labor market programs for the unemployed. For more detail, see the institutional description and literature review in Section 10 of McCall et al. (2016).

[8] A chunk of the non-response reflects individuals exercising the option that Danes possess to opt out of all research-related surveys.

We focus here on the findings in KSS related to formal training, which in the AES context means training intended to lead to a formal qualification. Pooling the two surveys, among those with a spell of formal training in the administrative data that lies within the relevant survey time window, about 77 percent self-report a formal training spell. On the other side, about four percent of those without a spell of formal training in the administrative data self-report one on the survey. The qualitative pattern of substantial under-reporting by some respondents combined with modest (though not trivial in a proportional sense) over-reporting by others echoes findings from other studies.

KSS dig a bit deeper by drawing on the theoretical framework provided in Bound et al. (2001). In regard to formal training, they find evidence consistent with "telescoping," a fancy way of saying that respondents sometimes remember incidence but have the timing a bit off. Expanding the window in the administrative data to include three months prior to the nominal time window in the survey question removes a fair amount of under-reporting. They also find that individuals who dropped out of courses or failed an exam at the end have much higher probabilities of not reporting their spell in the survey. This finding has important implications for the extent of bias in impact estimates of training, as these individuals presumably have smaller treatment effects than those who successfully complete their courses. Finally, they find a clear pattern related to salience, as very short courses induce more under-reporting than longer ones, another finding with real implications for impact analyses like those we review in Section 5.8. These findings illustrate the value of having not just multiple reports of training (e.g. one from the employee and one from the employer) but of having detailed, accurate administrative data on course timing, characteristics and completion as well as incidence.

3.4 Indirect evidence

3.4.1 Frazis and Lowenstein

Frazis and Loewenstein (2003) propose a new econometric approach to estimating linear models with mismeasured binary independent variables. To illustrate the application of their method, they study the effects of self-reported workplace training on wages using data from the 1987-1994 waves of the NLSY-79. Their data contain only a single measure of training, but their approach implies bounds on the extent of both false positives (non-trainees reporting training) and false negatives (trainees not reporting training).

They obtain a relatively tight upper bound of around 0.02 on the fraction of false positives, though even 0.02 matters in a proportional sense relative to an overall self-reported training participation rate of 0.13. In contrast, their procedure yields quite wide upper bounds around 0.70 on the fraction of false negatives. The upper bound on false positives matches our prior and the upper bound on false negatives contains little information. Reading their paper only reinforced our view that measurement error wreaks havoc on estimates and so is best avoided if at all possible.[9]

3.4.2 Smith and Whalley

The (very) small literature that seeks to validate survey measures of training provided by governments as part of workforce (i.e. active labor market) programs sheds some indirect light on the problem of measuring firm training. We know of only one unpublished study, Smith and Whalley (2006)[10][11] They make use of survey and administrative data on service receipt collected as part of the experimental evaluation

[9]Frazis and Loewenstein's bounds require instruments. They use gross job creation and destruction rates at the two-digit industry level and years of schooling for instruments in their wage growth equation; see the discussion on page 166 of their paper for their justification. We expect the instrument police have concerns.

[10]Their study consists, for the moment, solely of tables, figures and slides from a conference presentation.

[11]Perhaps surprisingly, the recent U.S. WIA Gold Standard Evaluation, though it collects both survey and administrative data on classroom training and other service types, does not do comparisons of this sort. Instead, it implicitly treats the survey responses as truth. See Fortson et al. (2017) and McConnell et al. (2021)

of the U.S. Job Training Partnership Act (JTPA) Program.[12] Measured by persons served, JTPA was the largest U.S. federal training program for most of the 1980s and 1990s.[13]

JTPA provided several service types. Classroom Training in Occupational Skills (CT-OS) has the most relevance for the topic at hand. It resembles off-site worker training in that it takes place in a classroom at a provider operating under contract to the program, often a local community college. It differs (on average) from off-site formal training paid for by firms in having a much longer typical intended duration (i.e., months rather than days or weeks). As described earlier, we might expect the longer duration to make the training more salient and thus increase recall rates. JTPA also provided subsidized on-the-job training, job search assistance, and adult basic education. The latter two have no direct analogue in the world of firm training. Subsidized on-the-job training in JTPA meant financial encouragement for a firm to hire a worker they might not otherwise; those workers then typically received the same training as other new hires.

Smith and Whalley compare the survey responses of treatment group members to the corresponding JTPA administrative records.[14] The administrative data measure only services paid for by the JTPA program. The service types provided in the administrative data comprise classroom training in occupational skills (CT-OS), subsidized on-the-job training at private firms (OJT), job search assistance (JSA), adult basic education (BE), work experience (WE), and "other".

In contrast to the administrative data's focus on JTPA services, the survey asks about education and training experiences more generally, with no specific mention of JTPA. This allows the survey to provide a consistent measure of education and training received for the experimental treatment and control groups. The survey respondent filled in a calendar with all spells of employment and all spells of education or training

[12]Bloom et al. (1997) and Orr et al. (1996) describe the evaluation in detail.

[13]Barnow and Smith (2016) provide extensive detail on JTPA and the other programs in the series and survey the related evaluation literature.

[14]The JTPA study yielded two separate administrative measures of services received. Smith and Whalley (2006) used the one produced by Abt Associates.

during the relevant reference period. After a long series of questions about the employment and non-employment spells, the interviewer then asked, for up to three spells of school or training, the start and end dates, the type (with 16 choices), average hours per week, whether or not the family paid at all and, if they did, how much they paid.

Smith and Whalley (2006) face the issue of how to map the 16 types of education and training captured by the survey into the six service types represented in the administrative data. For the classroom training and on-the-job training that interest us here, they code the types "high school", "two year college", "four year college", "graduate school", "business school", "technical institute", and "other vocational" as classroom training in occupational skills and the type "on-the-job" training as on-the-job training. This mapping seems reasonable to us, but represents another source of potential measurement differences between the two data sources.

More generally, in the JTPA Study context, neither the survey nor the administrative data represents "truth" both because they seek to measure somewhat different things and because of institutional features such as the failure of the survey designers to mimic the categories available in the administrative records and the fact that some treatment group members received light touch services prior to enrollment, services that they may recall on a survey but which do not necessarily appear in the administrative data.[15] Heckman et al. (2000) show using these data that a substantial fraction of the experimental control group received services similar to those provided to treatment group members by JTPA. Perhaps surprisingly, some treatment group members receive services outside of JTPA rather than JTPA services. These individuals will correctly report receiving training in the survey but not have a record of JTPA training in the administrative data.

Smith and Whalley (2006) offer several relevant findings. First, surveys miss a lot of service receipt. For example, of those with a spell of classroom training (defined as either CT-OS or BE) in the administrative data, only about 56 percent report

[15] This phenomenon resulted from the incentives implicit in JTPA's performance management system. See Kemple et al. (1993) for discussion.

receiving classroom training in the survey data, indicating 0.44 “false negative” rate that is distressingly close to Frazis and Loewenstein’s 0.70 upper bound on the false negative rate we discussed above. Second, both unconditionally and conditional on site indicators and demographics, survey respondents do better at reporting services that happen in classrooms, mainly CT-OS but also BE, than other services such as OJT. Third, some respondents have trouble distinguishing between certain pairs of services, especially CT-OS and BE and, more surprisingly, CT-OS and OJT. Overall, the results in Smith and Whalley (2006) for workforce services provided via a government program qualitatively resemble the findings in the literature on measuring firm training reviewed in the preceding subsection.

3.5 Some (mis)measured conclusions

The most obvious conclusion from our review of the small literature on training measurement is that survey measures of firm training, whether it takes place on-site or off-site, in a classroom or at the worker’s desk or station, contain large amounts of measurement error. That measurement error goes almost entirely in one direction, namely that of under-reporting of training.

To illuminate why measurement error matters in this context, consider first a binary measure of training incidence and suppose that the researcher seeks to estimate the causal effect of training incidence on some outcome. Let T denote the observed binary measure of training incidence, T^* the true binary measure of training incidence, and m_T the measurement error resulting from a difference between the two. In equation form, we have:

$$T = T^* + m_T \tag{1}$$

In the binary case, $m_T \in \{-1, 0, 1\}$. Furthermore, as noted at least as early as Aigner (1973), $\text{cov}(T^*, m_T) < 0$. Imagine further that our researcher wants to start by assuming conditional independence and a common effect linear world, and so estimates an equation like:

$$Y_i = \alpha_0 + \alpha_T T_i + \alpha_X X_i + u_i \tag{2}$$

where Y_i denotes the outcome of interest and X_i a vector of conditioning variables thought to suffice for conditional independence.

Estimation of (2) by ordinary least squares in the presence of measurement error in T_i implies attenuation of $\hat{\alpha_T}$ toward zero relative to the common effect α_T under relatively weak conditions. With measurement error in T_i, some observations coded as trained (i.e. with $T_i = 1$) did not actually receive training while some observations coded as not trained actually received training. Both directions of measurement error reduce the strength of the treatment contrast via contamination, keeping in mind that $\hat{\alpha}_T$ equals the conditional mean difference in the outcome Y_i between those with $T_i = 1$ and those with $T_i = 0$. Things become more complicated for a couple of reasons in a world of heterogeneous coefficients α_{Ti}, especially if the value of the coefficient correlates with the nature and incidence of the measurement error.

Now suppose that our clever researcher remembers the role of instruments in dealing with measurement error, and that our researcher just happens to have a relevant instrument z at hand for this purpose. As noted by Bound et al. (2001), Frazis and Loewenstein (2003), and several others, IV estimation will lead to anti-attenuation bias (i.e. bias away from zero) in the $\hat{\alpha}_T$. The attenuation bias from the measurement error, and the anti-attenuation bias from using an instrument to try to solve the measurement error problem when the measurement error is non-classical, combine to yield a net bias that depends on features of the particular empirical context.

A similar line of reasoning applies to the case of continuous measures of training. Consider such a measure, and call it H for hours of training. Suppose our mythical researcher wants to estimate a version of (2) with the continuous hours measure replacing the binary training indicator:

$$Y_i = \gamma_0 + \gamma_H H_i + \gamma_X X_i + v_i \tag{3}$$

To keep the discussion simple, we remain in a common coefficient world and consider measurement error of a particularly straightforward form, namely that either the respondent reports the correct value of training hours or they report zero hours. In notation:

$$H = H^* + m_H \tag{4}$$

with $m_H \in \{0, -T^*\}$. This simple case implies that $\text{cov}(H^*, m_H) < 0$, a result that need not hold in general but that will likely hold in the presence of strong asymmetric measurement error of the sort found in Smith and Whalley (2006) for training incidence.

Some simple algebra provides a general formula for the inconsistency in the instrumental variables estimator with a continuous training measure and an instrument correlated with the measurement error in that covariate:

$$\text{plim}\hat{\gamma}_H^{IV} = \frac{\gamma_H \text{cov}(H^*, z)}{\text{cov}(H^*, z) + \text{cov}(m_H, z)} = \gamma_H \frac{1}{1 + \frac{\text{cov}(m_H, z)}{\text{cov}(H^*, z)}} > \gamma_H \tag{5}$$

where the final inequality assumes that $\gamma_H > 0$. In other words, a negative correlation between the instrument and the measurement error implies bias away from zero in the coefficient of interest. As usual, complications arise with heterogeneous coefficients γ_{Hi}.

To illustrate the potential magnitude of the bias, we can construct an incidence measure from the JTPA data of Smith and Whalley and apply the bias formula of Frazis and Loewenstein (2003). The false negative rate of 0.44 reported by Smith and Whalley (2006) implies that the IV coefficient in a version of (3) without covariates would have an upward bias of 79% even in the absence of any false positives.[16]

The apparent empirical prevalence of under-reporting of training incidence (and,

[16] Frazis and Loewenstein (2003) show that the inflation factor for the anti-attenuation bias equals $1/(1 - \text{Pr(false negative)} - \text{Pr(false positive)}) = 1/0.56 = 1.7857$.

by extension, training hours and other intensity measures) has strong implications for the interpretation of the estimated effects of training in studies that rely on survey measures of training. In studies that compare the outcomes of self-reported trainees to those of self-reported non-trainees, the latter group will actually include many trainees, resulting in an understatement of the average effects of training to the extent that it actually affects outcomes. Of course, one might also wonder whether training reporting correlates with the impact of training. A natural conjecture would say that respondents better report training that has an impact on them. Alternatively, one may worry that the degree of recall bias is correlated with measured or unmeasured cognitive abilities, a concern that Frazis and Loewenstein (2003) raise. Alas, the literature allows little more than wonder as it has not yet, to our knowledge, addressed this important question in the context of firm training.

4 Evidence on the nature, extent, and determinants of firm training

4.1 Introduction

The purpose of this section is two-fold. We first present stylized facts about training participation drawn from a review of the existing literature. Beyond having value in their own right, these facts provide a valuable guide to the selection process into training for researchers doing impact studies. We split our presentation of previous studies into Section 4.2, which reviews the North American evidence, and Section 4.3, which reviews the European evidence.

Table 2 describes the surveys considered in our review. Within each row, the first column indicates the survey name, the second column two indicates the country or country group (i.e. OECD or EU) subjected to the survey, and the third column indicates the years in which the survey was administered. The latter equals a range or a set rather than a single year for panel surveys like the National Longitudinal Survey of Youth 1979 cohort (NLSY-79) in the United States and for repeated cross sectional surveys, such as the AETS in Canada. The fourth column describes the population surveyed in general terms and provides a rough sample size. The fifth column offers a terse description of the types of training defined and asked about in each survey. We dissect these definitional differences when we consider specific papers more deeply in Section 4. Finally, the sixth and seventh columns report basic estimates of training incidence and average training duration among the trained for the specific survey, as calculated in the source cited in the eighth column. We intend the estimates in this table to provide a high altitude view of what the literature offers in terms of incidence rates and mean duration.

In part because we continued to find additional studies as we worked on the chapter, we lack a formal rule for study inclusion. Our choices emphasize studies primarily devoted to studying participation in training, coverage of the most important datasets used in the firm training literature, our subjective assessments of study quality, and more recent studies. We recognize that we did not include a number of valuable, high-quality studies.

Like the chapter as a whole, our summaries of the various papers, and the stylized facts we glean from those summaries, focus in particular on issues of measurement, including survey response rates, as well as on the quantitative measures of training incidence and duration and their correlations with observed features of workers, firms and contexts. An interesting exercise that we do not attempt here would meta-analyze the broader universe of training studies, with the aim of decomposing the variation in unconditional training rates into components due to the population studied, the type of training, details of the survey measure of training, and so on.

[INSERT TABLE 2 HERE]

In addition to reviewing the extant literature, we provide in Section 4.4 a new study of participation in the four different modes of training described in Section 2. We do this not merely to "make the rubble bounce"; rather, we add value to the existing literature by presenting stylized facts about patterns of training across a multitude of countries using the same survey data source, namely the first three rounds of the PIAAC described in Section 3. As that section makes plain, and the reviews that follow illustrate, survey and administrative measures of training differ along multiple dimensions, including the populations surveyed (e.g. recently hired workers as in EOPP and SBA, a random sample of youth ages 14-21 in 1979 in the NLSY, all working-age adults, all workers, etc.), the reference period over which respondents report their training (e.g. since hired, during first three months of employment, during last 12 months, during last three years, during last 10 years, etc.), their exact definitions of the training they seek to measure, the path of gateway questions that lead respondents to the questions about training (or not), and in whether employers or employees respond to the survey.

On top of these differences in survey design and administration come the idiosyncratic sample selection and coding differences imposed by different authors using the same underlying survey data, differences that often lead to different magnitudes and patterns of training take-up. By providing a rich set of unconditional estimates of training incidence along with estimates of multivariate models of the determinants of different kinds of training in different countries using a common data set that covers the same population in each country and common specifications, we allow for cross-country comparisons well beyond anything available in the existing literature. Moreover, because the PIAAC collects information on all four modes of human capital accumulation presented in Section 2, it allows us to compare across training types and to examine the correlations between them.

4.2 Evidence from North America

This section summarizes the information on training participation in the United States and Canada provided by studies that use the related surveys listed in Table 2. We divide our discussion into subsections describing evidence from household surveys, employer surveys, and matched employer-employee surveys. One common theme unites the underlying studies: their age. Every study is at least 20 years old and some rely on data more than 30 years old. We expect that changes in the industrial mix of the North American economy (e.g. from manufacturing to services) as well as transformational changes in information technology imply different patterns in the present day.

4.2.1 Evidence from household surveys

Loewenstein and Spletzer (1999a) focus on the 1993 and 1994 rounds of NLSY-79, at which time the cohort was between the ages of 28 and 37. As noted above, two features of the NLSY-79 survey design in these waves affect the interpretation of their estimates. Respondents only reach the questions about informal training if they respond affirmatively to a question about the need for new job skills in the past 12 months. Loewenstein and Spletzer (1999a) note (in their understated way) that "the routing pattern in the survey ... limits the usefulness of the new NLSY informal training data for estimating the incidence and duration of informal training." The instrument imposes a slightly less restrictive condition on formal training through the last clause in the question: "Since [date of the last interview], did you attend any training program or any on-the-job training designed to help people find a job, improve job skills, or learn a new job?" In both cases, the survey design implies that the analysis provides a lower bound on training incidence.

Loewenstein and Spletzer (1999a) aggregate the underlying training types differently than we do in Section 2 or than the recent surveys from the OECD, Statistics Canada, or Eurostat do. In particular, they define a broader notion of formal training than we do that includes "apprenticeship, business school, vocational technical school" (1.5%), "on-the-job training" (7.6%), "inside seminars" (3.3%), "outside seminars" (3.8%) and "other training". Similarly, their measure of informal training includes "classes or seminars" (10%), "showed by ... supervisor (26%), coworkers (13%)", and "make use of any self-study material, ... manuals, workbooks, or computer-assisted teaching programs" (20%), some of which we would instead classify as casual training or even learning-by-doing.

Using their classification of formal and informal training, average annual incidence of formal training equals 17%. Informal training has a higher participation probability of 40%. Formal training participants average 48 hours, with a standard deviation of 127 hours. Informal training participants average 57 hours, with a standard deviation of 168 hours. In the NLSY-79, as in other surveys, mean hours of training disguise substantial heterogeneity in duration. Using our definitions of formal and informal training in Section 2 in place of theirs reduces the estimated take-up of formal training to well below 10% (how much depends on whether we consider "outside seminars" formal or informal) and reduces the incidence of informal training by more than half.

Loewenstein and Spletzer (1999a) find that formal and informal training are strong complements. Having participated in informal training bumps up the conditional probability of participating in formal training by 0.12 (evaluated at the mean of their covariates), a near doubling of the conditional probability. Formal and informal training both exhibit strong positive correlations with AFQT scores[17] and years of schooling. A worker with less than 12 years of schooling has a 0.08 lower conditional probability of participating in formal training compared to one with exactly 12 (remember the mean is 0.17!) while someone with 16 years has one 0.05 higher. Female workers have a 0.03 higher conditional probability of participation than male workers, and Hispanic workers have a 0.04 higher conditional probability of participation in informal training than do white workers. Black workers do not differ statistically or substantively from white workers in their conditional probability of either type of training. Working for a firm with multiple sites increases the conditional probability of participation in both types of training, as does the log of firm size. The conditional training probabilities appear concave in tenure (up to eight or nine years). While we have (much) more to say about training and tenure in Section 5, the NLSY-79's clear picture of ongoing, rather than front-loaded, training merits mention here.

[17] The AFQT (Armed Forces Qualifying Test) provides a measure of "ability" in the loose sense that labor economists use that term. Most NLSY-79 sample members completed the test. See, e.g. Neal and Johnson (1996) for more details and more about the meaning of the test.

Hui and Smith (2002) use the 1998 Adult Education and Training Survey (AETS) to study training in Canada. They focus on individuals 25 to 64 years old who are not full-time students at the time of the interview. As noted in Table 2, the AETS was an occasional supplement to the Canadian Labour Force Survey (LFS) between 1983 and 2002. The 1998 AETS had a response rate of 81% which leaves the authors with a sample of more than 23,000 respondents. The authors discard all training that is not work-related[18] and divide activities into what they label "training programs" and "training courses" where "programs" lead to formal certification and "courses" do not. As such, their categories roughly correspond to the formal and informal categories defined in Sections 2 and 3 and we label them as such going forward. Finally, the authors partition the two types of work-related training into three bins based on their financing: employer-financed (in whole or in part), self-financed (in whole or in part, with no employer financing), and financed by the government (with no employer or self-financing).

Over the 12-month recall period, seven percent of the respondents report participating in formal training and 22 percent in informal training. Employers finance about 35 percent of the formal training and a bit over 70 percent of the informal training. Conditional on participation, formal training averages 363 hours whereas informal training averages 39 hours. These averages (again!) miss substantial heterogeneity in hours, with standard deviations of 501 and 100 hours for formal and informal training, respectively. Both types of training exhibit strong bivariate correlations with initial schooling levels that largely disappear when conditioning on job characteristics and demographics. The authors interpret the lack of a (conditional) correlation as evidence against Canadian workers facing liquidity constraints when making investments in their human capital.[19]

[18]The OECD PIAAC and the European AES surveys collect data on all training episodes and then ask respondents to distinguish between work-related and recreational training. In contrast, the Canadian AETS collects information about work-related and recreational training in separate survey modules.

[19]See Section 6.1.2 for more on credit constraints.

4.2.2 Evidence from employer surveys

Barron et al. (1997b) and Black et al. (1999) study training participation among recent hires in both the 1982 Employment Opportunity Pilot Project (EOPP) and the 1992 Small Business Administration (SBA) datasets. The 1982 follow-up survey of EOPP obtained responses from 70% of the 5,700 employers that completed the first survey in 1980, but only 1,916 (or 34%) of the establishments survive the authors' list-wise deletion (for missing covariate values) in the subsequent analysis. The 1992 SBA survey was administered to 3,600 establishments over the summer of 1992. The SBA survey over-samples large firms, but not (unlike EOPP) low-income workers.[20] 1,123 (31%) establishments completed enough of the survey to make them useful. The authors apply sample weights to both datasets to account for their (different) stratified sampling.

Both surveys contain four measures of training: on-site formal training, informal training by a supervisor, informal training by a coworker, and a measure of time spent watching others perform tasks. The SBA survey adds off-site formal training to the set. All these training measures cover the first three months of employment for the most recent hire. Just as in the NLSY-79 study by Loewenstein and Spletzer (1999a), these measure do not fall neatly into our four training types from Section 2. It is not obvious, for example, how structured the management or co-worker training is, and thus whether we would categorize it as informal training or as casual training.

The two surveys agree that virtually all (0.96) new hires receive some training during their first three months on the job. Training time over the same period averages about 150 hours in both datasets, keeping in mind that initial worker training may last longer than three months in some jobs. Firms supply formal training to 12% of newly hired workers in the EOPP data and 14% in the SBA data. Around 90% of new hires receive informal training by management and 60% receive informal training by their co-workers in both surveys. The two surveys disagree somewhat regarding the extent of training

[20] Only 10% of the newly hired in the (unweighted) EOPP sample had a college degree whereas 25% of the (unweighted) SBA sample had a college degree. The newly hired in both surveys were on average slightly younger than 30. Unweighted mean establishment size was 72 for EOPP but 182 for SBA.

via "watching others", with an incidence rates of 0.80 in the EOPP data and 0.65 in the SBA data.

Barron et al. (1997b) provide some additional statistics based just on the SBA data. Conditional on participation, formal training averages 66 hours over the first three months on the job while workers spend an average of 68 hours watching other workers perform tasks. The total number of hours of training during the first three months vary widely by level of education, with high school graduates (in 1992) receiving an average of 150 hours of training but college graduates receiving 70% more than this. Almost as stark is the difference across firm size: Workers at firms with fewer than 25 employees average around 150 hours of training but workers at firms with at least 250 workers average as many as 220 hours. The employer respondents indicate no differences in hours of training by worker sex or race (a finding that also appears in the EOPP data).

The authors run logit models for the types of training measured in each dataset for the subsample of respondents without missing values for any the explanatory variables used (1,471 for EOPP and 888 for the SBA, or around 25% in both cases). They find firms more likely to offer off-site training to college graduates and union members and larger firms more likely to send their workers off-site. College graduates in both data sets have higher conditional probabilities of receiving on-site formal training, as do union members in the SBA data. Larger firms also have a higher conditional probability of offering formal on-site training to their workers according to both surveys. Fewer differences emerge when the authors study informal training, probably because (nearly) every new hire gets at least some of this. Larger firms are more likely to offer all three types of informal training but other than that, not much shows up in terms of statistically or economically meaningful differences.

Lynch and Black (1998) study participation in firm training using the 1994 Educational Quality of the Workforce National Employers Survey (EQW-NES), a nationally representative survey of private establishments with more than 20 employees.[21] The

[21]The exclusion of firms with few workers results in a sample representative of 75% of private sector

EQW-NES surveyed plant or local business site managers by telephone in August and September 1994. Of the 4,633 establishments contacted by the EQW-NES, 2,945 (64%) completed the survey. Very large establishments were marginally less likely to participate. The authors weight their data when relevant.

The EQW-NES defines formal training to "[include] all scheduled training events that had a predefined objective" a definition broader than the one we adopt in Section 2. Its definition encompasses training offered on-site and off-site, and during and not during regular working hours. The underlying survey question simply asks "Does your establishment pay for or provide any structured or formal training either on the job (by supervisors or outside contractors) or at a school or technical institute?" The structure of this question complicates comparisons with other studies for (at least) two reasons: First, the question refers to an establishment rather than individual workers; put differently, a correct positive response does not imply that all workers in the establishment receive the training. The question also offers no time frame, so a correct positive response could imply, for example, no training in the current year but training before that and/or in the future.

With those caveats lodged firmly in our consciousness, this formulation yields an incidence rate of 81%; establishments with fewer than 100 workers have a rate of 75% while those with at least 1,000 workers have a rate close to 100%. The rate also varies across industries: for example, it equals 61% for textile and apparel but 100% in communication and finance. Establishments vary greatly in the types of training offered. Roughly a quarter of reporting establishments provided basic education such as literacy and numeracy skills, about half offered training in computer skills and teamwork, and around 80% offered safety training. These fractions in turn vary by industry. For example, few firms in business services, retail or construction offer basic education ($< 20\%$) whereas more than half of the establishments in utilities, finance, insurance, and primary metals do.

The EQW-NES also asked managers about casual training such as watching oth-

workers but only 15% of establishments.

ers doing a job. This had a reported incidence rate of 100% which necessitates its omission from the multivariate analyses. Finally, about half of the respondents offered a "guestimate" of their establishment's annual expenditures on formal training. This (potentially selected) sample reports spending around 5% of their total labor costs on training, roughly the same as they report spending on recruiting.

A further question asks managers "In the past year, how many workers received formal instruction?" Combining the response to this question with information on establishment size yields an annual formal training rate for each establishment. Taking an appropriately weighted average of these rates yields an EQW-NES estimate of the fraction of workers involved in formal training that exceeds 40%. This represents a considerably higher rate than, for example, the NLSY-79 estimate of 0.17 discussed above.

In addition to providing basic descriptive statistics, the authors study the determinants of training (any, as well as specific types) using firm characteristics (size, industry, multiple establishments), worker characteristics (educational level, fractions of newly hired, minority and female workers), and workplace characteristics (use of TQM, worker rotation, self-managing teams, number of organizational levels, benchmarking, and workers per manager).[22] Specifically, Lynch and Black (1998) estimate logit models for four of the 11 types of formal training: computer training (offered at 52% of establishments), teamwork training (53%), basic education (25%), and sales and customer service training (68%). They find that providing formal training correlates with establishment size, establishment changes in employment (both up and down), and workers' educational levels. None of the other included worker characteristics, including demographics and occupational composition, matter statistically or substantively for any of the training types. In contrast, benchmarking and TQM strongly increase the conditional probability of all four types of training.

Just a year after the EQW-NES, Canada conducted a similar survey, the Ekos

[22] Benchmarking means to "compare practices and performances with other organizations" according to the survey. Younger readers should note that "Total Quality Management" (TQM) swept the organizational world of the 1980s, leaving behind stacks of unread how-to guides and the occasional suggestion box.

Workplace Training Survey (WTS). The WTS sample had two components: The first consisted of 1,089 establishments successfully re-surveyed from among 2,543 establishments participating in a similar survey two years prior. The second was drawn from a Dun & Bradstreet file of 7,300 establishments of which 1,495 were still in business, not single-person establishments, had a phone number in service, and proved willing to participate in the survey. Telephone interviews of the 2,584 WTS establishments took place in May and June 1995. Our summary relies on Betcherman et al. (1997), who weight the survey data to reflect the overall population of Canadian firms at the time.

The survey posed the following training question: “Has your <firm or establishment> sponsored or provided any formal or informal training to employees over the past year?” If the respondent answered yes to this question, a second question asked about the formality of the training. The WTS defines formal training as having predefined objectives and curriculum as well as a structured format, with the respondent given examples such as “classroom instruction, scheduled and structured on-the-job training, apprenticeship training, and courses at formal educational or training institutions that are paid for by the employer.” This definition more-or-less combines formal and informal training as defined in Section 2.

Seventy percent of establishments responded positively to the first question, but only 42% reported providing formal training as the WTS defines it. That number lies well below the 80% found in the EQW-NES, a difference too large to ascribe solely to the presence of an explicit recall period in the survey question. As training was more likely at larger establishments, 89% of all Canadian employees worked at establishments offering some kind of training and 76% at establishments offering formal training. Of course, the same point we raised with the EQW-NES applies here: presumably not all workers at firms offering any training, or formal training, actually receive it.

The authors estimate multivariate models of training provision similar to those in Lynch and Black (1998). As in the EQW-NES, they find larger establishments and establishments within larger firms more likely to provide training. Unionized establishments, unlike their American counterparts in the EQW-NES, turn out more likely to supply both any training and formal training, though the difference attains statistical significance only for formal training, with a corresponding odds ratio of 1.43.

4.2.3 Evidence from matched employer-employee surveys

Holzer and Reaser (1999) investigate training as measured in the Multi-City Study (MCS) of Urban Inequality, a dataset collected between June 1992 and May 1994 in the Atlanta, Boston, Detroit, and Los Angeles metro areas. The survey team collected data by phone from the individuals responsible for hiring at 800 firms (a 67% response rate) in each metro area. They drew one-third of the firms from household surveys administered in the same four areas and two-thirds from regional employment directories. The authors weight their results to reflect each area's overall population of employees.[23]

The study focuses on "the worker hired into the most recently filled job at the establishment" but without a specific time interval like the three-month window in the EOPP and SBA surveys discussed in Section 4.2.2. Holzer and Reaser (1999) end up with a sample of 2,260 workers with a median tenure of two months and 90% having tenure less than a year.[24] The wide variation in tenure among respondents, and thus in the time period covered by their responses, complicates interpretation of the incidence rates and statistics regarding hours of training estimated from these data. For instance, two percent of employers report having supplied their most recent hire with over 1000 hours of training! The MCS defines formal training as "a formal program with a professional instructor" and informal training as "training from co-workers or supervisors". Roughly, then, the definition of formal training in the MCS combines what we call formal and informal training in Section 2, while its notion of informal training corresponds to what we call casual training.

With the caveat regarding heterogeneous reference periods close at hand, the MCS

[23] Although we locate this study in our section on matched employer-employee surveys, neither Holzer and Reaser (1999) nor Holzer (1996) make much use of the potential linkages.

[24] The sample drops to 1,847 for the analysis of informal training, presumably due to a higher level of item non-response.

data yield participation rates of 0.44 for formal training and 0.91 for informal training. The rate for formal training is nearly double that from the SBA and EOPP datasets. In contrast, the informal rate tells a story very similar to that told by the SBA and EOPP, namely that management expects that nearly all new hires receive some informal training. The unconditional rates vary by demographic group: Black women have a participation rate in formal training of 0.56, black men 0.48, white women 0.43, and white men and Hispanics 0.39. Conditional on participation, Hispanics spend on average 52 hours in formal training, black women 61 hours, black men 66 hours, white women 83 hours, and white men 116 hours. Average hours of training increase with education level, from a low of 42 hours in formal training for high school dropouts to a high of 107 hours for college graduates. Employers report no meaningful differences in informal training participation rates by demographic group or completed schooling. Like the other studies we have considered, the MCS data reveal tremendous heterogeneity in hours of training among those who receive it. For example, the standard deviations of training hours equal 84 for black women and 210 (nearly double the mean) for white men, and equal 115 hours for dropouts and 175 hours for college graduates. Looking across subgroups, the standard deviation typically increases with the mean.

The authors estimate a Tobit model of training hours with a lower limit at zero and an upper limit at 100 hours.[25] The model includes worker, job and firm characteristics. Nearly all of the demographic and educational differences so apparent in the descriptive statistics fade into statistical and substantive insignificance in the multivariate model. In regard to the educational differences, the authors note that firm size has strong positive correlations with both worker educational levels and hours of training.

Frazis et al. (2000) exploit the data from the 1995 Survey of Employer-Provided Training (SEPT-95). The SEPT95 captures both the incidence and intensity of training. It surveys both employers and employees, which makes it possible for the authors to study the consequences of omitting, say, firm characteristics, in multivariate models

[25]A modern reader might object that the zeros are real zeros and not censored negative values of training hours, and that training likely has a fixed cost aspect, both of which render the Tobit model potentially problematic.

of training incidence or intensity, that also include worker characteristics. Most notably, the SEPT95 had both its employer respondents and its employee respondents keep real-time logs of training activity, which should imply fewer issues with recall bias in the subsequent analyses.

The SEPT-95 drew a (stratified) sample of 1,433 establishments from among firms with more than 49 employees during the fourth quarter of 1993. 1,062 employers (74%) filled out their questionnaire but only 949 (66%) filled out their logs. Two employees were selected at each establishment at random. Of these, 1,074 employees (51%) filled out the questionnaire and 1,003 (47%) completed the real-time log. The interviews took place between May and October 1995.[26]

The authors define formal training events as those "provided or financed by the establishment", a definition both broad and vague and likely encompassing all of what we define as formal, informal, and casual training in Section 2. To obtain information on hours and costs of formal training, the BLS interviewers administered their survey to either "the training or human resource director [...] or the person who handled personnel and training issues." Cost items included money spent during calendar year 1994 on wages and salaries for in-house training, fees paid to outside training companies, and tuition reimbursements. The interviewers did not collect information on materials or overhead nor on the overall training budget (i.e. the total amount spent on training during the year).

The SEPT-95 asked employer respondents to keep a log during a two-week period to help them recall the number of employees receiving formal training, the hours and type of training received, and who conducted the training. On the other side of the employment relationship, the SEPT-95 first interviewed workers about their basic background characteristics like age and sex, along with characteristics of their employment, such as occupation, earnings, tenure, and training experience over the previous 12 months. The worker interview next involved the completion of a log covering all activities in

[26]The sample was stratified by industry and employer size category. The refusal of some employers to allow the BLS to contact their employees drives (in part) the lower unit response rate for employees than employers. This refusal negatively correlates with supplying formal training.

which they were "taught a skill or provided with new information to help them do their job better" for the three calendar days prior to the interview. For each activity, "questions were asked concerning who or what helped them learn the skill or information, how they learned the skill or information, what type of skill or information was learned, and how much time was spent learning this skill or information." In addition to providing three days worth of information on training activities, the log component of the worker survey provided training in how to complete the self-administered seven-calendar-day log of training events that the workers received following the interview. As with the employer logs, the response rate to the self-administered worker logs of 48% lies well below the corresponding survey response rate. Combining the survey log with the self-administered log provides 10 days of detailed information for a substantial subset of the workers in the study.[27]

At 93%, the fraction of establishments claiming on the employer survey to have provided some formal training to their employees some time during the previous 12 months substantially exceeds both the 81% reported by Lynch and Black (1998) and, perhaps more surprisingly, the 70% estimate based on the worker survey. That worker estimate in turn well exceeds the corresponding estimates of 17% for formal training and 40% for informal training in the NLSY-79 reported by Loewenstein and Spletzer (1999a). These SEPT-95 estimates derive not from the logs but from direct survey questions about training in the past 12 months; as such, the differences should not result from recall issues.

We observe much closer agreement between employers and employees in the SEPT-95 for the number of training hours per worker calculated using the training logs in both cases. Re-scaled to cover the period from May to October 1995, workers spent on average 11 hours in formal training according to their employers and 13 hours according to themselves.[28] The mean hours measure for informal training, reported only by the

[27]The BLS applied an algorithm to the responses to the "who or what" and "how" questions to classify each activity as formal training, informal training, or self-learning. The authors do not analyze the self-learning material and do not count it in their measures of training time or incidence.

[28]Sadly, the authors do not report a correlation of the two measures despite the title of their paper.

workers, equals 31, for a total of 45 hours of training. Assuming a 40-hour work week and no seasonal patterns implies that training consumes four percent of all work hours.

Wages paid to workers while in training (averages of $225 and $420 during formal and informal training, respectively, both in 1995 dollars for the period May to October 1995) represent the largest components of the overall cost of training to firms. For calendar year 1994, firms report smaller (but not trivial) average expenditures of $140 on in-house instructors, $100 on outside instructors, and $50 for tuition reimbursements for formal training activities. Re-scaling, summing, and price-adjusting these values using the CPI yields annual training expenditures in 2022 dollars of around $3,050 per worker, for a population of workers with an average age of 40 and close to five years of tenure. By comparison, Hanson (2022) estimates average in-state tuition at public colleges in the United States at about $9,350. Using the SEPT-95 estimates, the average firm spends this much on training a worker every three years. Put differently, every 12 years firms on average spend as much on training as a worker would spend obtaining a four-year degree at the local public college.[29]

We round out this subsection with a discussion of the participation patterns coming out of the Canadian Workplace and Employee Survey (WES). The target population of WES consisted of all business locations operating in Canada with paid employees, with the exception of agriculture, religious organizations, and public administration. Statistics Canada fielded this longitudinal study every year from 1999 to 2006, with top-ups of new business locations every two years. Making the survey mandatory for businesses (but not workers) led to a very high response rate of 0.95, with non-responding locations either out of business or seasonally inactive. Participating employers provided lists of employees to the interviewer; Statistics Canada would attempt to interview a maximum of 24 of them, chosen at random, with each (ideally) remaining in the panel for two years. The employee part of the WES realized a response rate of 0.83 in 1999, resulting in a sample of 23,540 employees at 6,322 business locations. Turcotte

[29] Of course, this comparison ignores the time cost of going to college but includes the time cost of the firm training. It also ignores the subsides that keep public college tuition well below social cost.

et al. (2003) study the first year of the WES. They weight their results to account for stratified (by industry, region, and size) sampling and panel attrition (i.e. “stratum erosion”).

The employer part of the WES first collects information about workforce size and composition, along with compensation practices and benefits before asking about the nature and extent of workplace training. The survey asks the employer respondent to “include all types of training intended to develop your employees’ skills and/or knowledge through a structured format, whether it takes place inside or outside the location.” The first set of training questions relate to “classroom training”, which the survey defines for the respondent as “all training activities which have a pre-determined format, including a pre-defined objective, a specific content” and where “progress may be monitored and/or evaluated.” After a series of questions about the use and financing of classroom training, the survey continues on with the collection of information about OJT activities. No framing is offered in this part of the survey. Instead, the respondent is presented with a series of examples to select from, such as “orientation for new employees”, “managerial/supervisory training”, “team building, leadership, communication”, and “occupational health and safety”.

About 31% of interviewed locations report providing some classroom training while 45% report providing some OJT. Both rates increase in business location size and evince considerable heterogeneity across industry. Rates of providing classroom training range from a low of 19% in real estate to a high of 59% in finance and insurance, while rates of providing OJT range from 27% in real estate up to 64% in finance and insurance. Conditional on providing some training, the data reveal surprisingly little variation with either training type or firm size in the proportion of the workforce receiving training. It equals equals 63% for classroom training and 66% for OJT. Two out of three workers participate in both types of training in firms with fewer than 20 employees compared to around 50% in firms with more than 100 employees.[30]

[30]The survey also collected information about training expenditures but the authors unfortunately do not make use of it.

The employees selected for interviews were instructed to think about "job-related training provided or paid [for] by [their] employer" and then asked the same questions about classroom and on-the-job training. Surprisingly, the former (37% of business locations) appears slightly more prevalent than the latter (30% of business locations). As in previous studies, both increase in educational attainment. In contrast, while classroom training increases with tenure, on-the-job training decreases. The authors find no evidence of differences in training take-up between men and women but do find strong provincial differences, with the probability of classroom training much higher in Quebec than in the other provinces, a point we return to in Section 6.2.2.

As the locations offering classroom training represent two-thirds of total employment and those offering OJT represent three-quarters of it, the employer survey responses imply that just over 40% of employees in Canada participate in classroom training and around 50% participate in OJT. The rate for formal training basically agrees with the rate from the employee survey given in the preceding paragraph, but the OJT rate differs by a quite substantial 20 percentage points. We do not have a good explanation for the difference, and Turcotte et al. (2003) do not provide one either. Moreover, the scheme for sampling employees in the WES (i.e. at random rather than just the newly hired), quite distinct from that in e.g. BBB (1997a,b), makes it challenging to sort among competing explanations.

4.2.4 Summing up

The evidence from the (mostly rather dated) North American literature on training displays remarkable disagreement on basic descriptive statistics related to training. For instance, less than 20% of workers report receiving formal training over a 12-month period according to the NLSY-79 compared to more than 70% in the SEPT-95. Some of this heterogeneity in estimates of training incidence flows out of the heterogeneity in the definitions of formal training that we remarked on in Section 3. For example, Loewenstein and Spletzer's (1999a) measure of formal training covers everything from business school classes to generic "other training" whereas Hui and Smith (2002) classify training based on whether or not completion yields a formal certificate. Such differences make it nearly impossible, given the lack of within-study comparisons of different measures in the extant literature, to sort differences of substance from differences of measurement.

On the bright side, all of the surveys considered in this subsection agree on two big-picture conclusions: First, training, whether formal or informal, is ubiquitous. Second, firms spend substantial resources on training, with estimates around four to five percent of labor costs according to EQW-NES and SEPT-95. We return in Section 4.4 to the question of whether or not the conclusions of these older studies remain valid for the more recent U.S. labor market and, more broadly, for some other current OECD labor markets. Put differently, we investigate the temporal and spatial external validity of the big picture findings from the older North American literature. We can think of plenty or reasons to doubt that external validity, particularly over time, given changes in rates of college going, China's entrance into the WTO in 2001, changes in production technology and sectoral mix, and developments in IT.

4.3 Evidence from Europe

We now turn to the European evidence. As noted in the introduction, although apprenticeships clearly constitute firm training, and despite their institutional importance in the European context, we omit papers solely focused upon them from our review. We do so both because the literature already contains two fine reviews by Wolter and Ryan (2011) and Carruthers and Jepsen (2021) and because our chapter already contains many pages. We organize the European review as we did the North American one, by first considering studies based on household surveys and then studies based on employer surveys. Unfortunately, we did not find any European studies making use of matched worker-firm data.

4.3.1 Evidence from household surveys

Pischke (2001) uses the German Socioeconomic Panel (SOEP) to study training for the years 1986-1989.[31] As best we can tell, it is the first study of training using the SOEP data and published in an English-language journal. The SOEP, a German analogue to the PSID, comprises a representative sample of 4,500 households, each surveyed annually since 1984. Its low initial response rate of 61% resulted from including only households in which all household members 16 years or older successfully completed a baseline interview. The 1989 wave of the SOEP added a module on training; the responses from that module form the basis of the study. The 1989 wave achieved a re-interview response rate of 92% of the originally responding West German households; the SOEP added a sample drawn from the eastern Länder the following year. Pischke (2001) weights his analyses to account for both stratified sampling and non-response.

The SOEP posed the training questions to all respondents ages 16-64, no matter their labor market status. For his analysis of training incidence Pischke includes only respondents employed as of the 1986 interview. The 1986 interview falls at the start of the 3-year reference period for the SOEP training questions that we noted in Table 1. This restriction should remove most training received while unemployed from the analysis. Unlike most of the other training studies we consider, Pischke (2001) retains younger respondents, public sector workers, and the self-employed. The broader population of interest implies extra care in comparisons with other studies. Caveats held tightly in hand, the study finds that about 28% of the 3,413 respondents report participating in at least one training episode over the three years prior to the 1989 interview, and half of those who participate recall at least three such events. Roughly two out of three training events last less than a week, and roughly six in ten trainees receive a certificate. Although, as the author notes, about three-quarters of the certificates only certify participation, they still make the training verifiable to outsiders. Despite this, 83% of the training incidents take place during work hours, organized or at least partly funded by the employer.

[31] The "Sozio-oekonomischen Panel" auf Deutsch.

Upon estimating a linear probability model of any training participation on a "standard" set of regressors (i.e. sex, potential experience, years of schooling, immigrant status, part-time status, occupation type indicators, sector indicators, and indicators for firm size categories), Pischke (2001) finds that public sector workers have substantially higher conditional probabilities of training participation, as do the self-employed (0.14), white-collar workers (0.22), managers (0.34), and those in larger (2,000 or more) firms. The last three findings parallel those in much of the North American literature. Pischke (2001) also finds women conditionally less likely (-0.07) to receive training.

Leuven and Oosterbeek (1999) use data from the first (1994) wave of the International Adult Literacy Survey (IALS) to study training participation in Canada, the Netherlands, Switzerland, and the United States.[32] The IALS acheived decent response rates of 69% in Canada, 45% in the Netherlands, 55% in Switzerland, and 60% in the United States. The authors use individual sampling weights for all of their analyses. Leuven and Oosterbeek (1999) include all respondents ages 16-65 (what the IALS thinks of as "adults"). They appear to restrict their analyses to the employed, but do not define the precise nature of the restriction.

The authors describe the construction of their training measure as follows:

> Whether a person participated in any work-related training is deduced from a combination of the following questions "Did you receive any training or education since August 1993?" and "What was the main reason you took this training or education?" (Respondents are only counted if they give "career or job-related purposes" as the main reason), and "Were you taking this training towards . . ." (where we did not count those courses leading to a formal education qualification)

The list of options in the last question is "a) a university degree/diploma/certificate; b) a college diploma/certificate; c) a trade-vocational diploma/certificate; d) an appren-

[32]The authors omit the other first wave countries, namely France, Germany, Ireland, Poland, and Sweden. The Swedes did not ask the detailed training questions; the authors have comparability concerns in the other cases.

ticeship certificate; e) an elementary or secondary school diploma; f) professional or career upgrading; and g) other." We conjecture that they interpret "formal education qualification" to mean options a) through e). In terms of the definitions in Section 2, we see their measure as consisting mainly of what we call "informal training" though perhaps with some bits of formal training and casual training added in at the margin. They do not provide a clear substantive justification for excluding the bulk of what we call formal training. As documented in Section 4.2 for the U.S., workers receive a substantial amount of formal training, so this exclusion matters.

In terms of the basic descriptive statistics, the participation rates vary from a low of 0.29 in Switzerland to a high of 0.40 in the U.S. Conditional on participation, workers average between 1.5 weeks (the US) and 3.5 weeks (the Netherlands) in training. Employers' financial support for the training varies between 73% (Switzerland) and 86% (the Netherlands) of the episodes, although this survey (like most of the others) does not attempt to elicit the share of training costs borne by the employer.

The authors estimate participation probits separately for each of the four countries. Like most studies published before their own, they find that training take-up increases with formal educational level. Like most studies of similar vintage that estimate probit or logit models, they report difficult-to-interpret coefficient estimates rather than average derivatives (i.e. mean marginal effects), which limits discussion of the substantive meaning of their estimates as well as comparisons with other studies. They find that women have a lower conditional probability of training in all four countries, but only in the Netherlands (at the five percent level) and the U.S. (at the 10 percent level) do the coefficients differ statistically from zero. The conditional association between age and their training measure varies substantially among the four countries. In Canada, it follows an inverse U-shape with 36–45-year-olds most likely to participate while in Switzerland it follows a U-shape with the same age group least likely to participate! In the Netherlands, training falls off at older ages while in the U.S. the youngest age group (i.e. 16-25 year-olds) has a much lower conditional probability than all of the others.[33]

Arulampalam et al. (2004a) study training participation across ten European countries—Austria, Belgium, Britain, Denmark, Finland, France, Ireland, Italy, Netherlands, and Spain—using the first six waves (1994 to 1999) of the European Community Household Panel (ECHP). They include respondents ages 25 to 54 who report working at least 15 hours per week. They exclude respondents working in agriculture and participating in apprenticeships and related training schemes. They further restrict their analysis sample to respondents present in at least two consecutive waves with valid values for all of the variables they include in their multivariate models, both the training variables and (more surprisingly) the conditioning variables.[34]

We mentioned the timing issues that arise with the ECHP due to failure of those fielding the survey to match the timing implicit in the question in Section 3. The authors note them too, and do what they can to avoid double-counting by using the (imperfect in some countries) reported training spell start dates. Their training measure builds solely on the ECHP survey question shown in Table 1, which concerns "vocational education or training, including any part-time or short courses?" They ignore responses to the parallel question, also included in Table 1, concerning "general or higher education" noting that very low rate of positive responses in their population of interest. They, and we, interpret their training measure as including all of what we call informal training as well as some of what we call formal training (i.e. short courses that lead to a credential but that respondents do not confuse with attending college more broadly). Given this interpretation, we might expect higher rates of training here than in, e.g., the Leuven and Oosterbeek (1999) study.

The descriptive statistics reveal considerable heterogeneity across countries, some

[33]The authors, as part of their broader purpose of considering the role of both workers and employers in training choices, note that "age as a proxy for the pay-off horizon matters for workers but not for the firm. This can be explained in part by the fact that training is a very imperfect statistic for expected tenure, which is the pay-off horizon for the firm."

[34]The paper does not indicate the fraction of the respondents lost to these various restrictions. Descriptive statistics are weighted to take account of "the survey design and patterns of individual non-response" but the multivariate analysis is not weighted, presumably because the statistical software used by the authors did not facilitate doing so.

fraction of which results from differences in the effective recall period across countries. On the upper end, more than 40% report participation in training in Britain and Denmark and one in three report participation in Finland. At the lower end, only 15% report participation in Austria, Belgium, and France, and less than ten percent report participation in Ireland, Italy, the Netherlands, and Spain. In nearly all of the countries the separate participation rates for men and women differ from the pooled rates just described by only 0.02 or less. Denmark, at 0.05, and Finland, at 0.07 constitute the exceptions. Only one country appears in both Leuven and Oosterbeek (1999) and Arulampalam et al. (2004a), namely the Netherlands. The data reveal a vast difference in training reports in this case, with this paper reporting a rate of 0.07 using the ECHP and Leuven and Oosterbeek (1999) reporting a rate of 0.32 using the IALS. As the datasets cover adjacent time periods, question wording effects head the list of suspects in explaining the difference.

For their multivariate analysis, the authors estimate a random effects probit with training participation in a given wave as the dependent variable. Their paper foregrounds differences by sex by estimating the multivariate models separately for male and female respondents, noting that the data reject the pooled model in nearly every country. The data further reject the null hypothesis of no random effects in every country, with the random effects typically consuming around 30% of the variation in the dependent variable. The models include age categories, indicators for fixed term contracts, seasonal work, part-time work, public sector employment, education levels, marriage and cohabitation, health issues, any children under 12, firm size categories (for the private sector), occupation, industry, and wage quintile, all measured in the preceding year, as well as year indicators.

The analysis reveals a strong negative age gradient in training incidence for men in every country, with varying intensity and nearly always statistically significant. In contrast, it reveals no strong relationship with age for women in any country. Male and female public sector workers consistently report more training, as do male and female respondents at higher educational levels. These findings parallel pooled results from many other studies. Part-time workers, fixed-term contract workers, and seasonal workers all tend to report less training, but the estimates for most countries turn out substantively small and not statistically distinguishable from zero. To our dismay, the authors do not present or discuss the results for the remaining covariates.

Booth and Bryan (2007) use waves 8, 9 and 10 of the British household panel survey (BHPS) to study training participation among Britons. The BHPS had an enviably high unit response rate of 90% for waves 8-10. It began asking training questions in its eighth wave and maintained them unchanged through the 18th and final wave in 2008.[35] The wording mimicked that of the NLSY-79 with the peculiar difference that the BHPS anchored the recall period in the BHPS to a particular date rather than defining a fixed window relative to the survey date. According to Buck et al. (2006), this led in practice to a marginally longer average recall period than in other similar surveys.

The authors limit their analysis to private-sector, full-time employees between the ages of 16 and 65 "with valid information on [their] main variables and who did not report more than a calendar year of training." They exclude person-years that include a spell of full-time education (2% of the sample) and drop 330 (or 11% of the sample) person-years with training but with missing type, place, duration, or financing information. They end up with 8,316 worker-year observations of which 31% include at least one training spell. The 31%, or 2,575 worker-years, covers a total of 5,272 distinct training spells. As the BHPS collected detailed information about the three longest training spells, their study includes detailed information for a subset of 4,317 training spells.

The participation rate of 0.31 exceeds that for the U.S. from the NLSY-79, despite an older population, but in line with a somewhat longer recall window and the inclusion of a small amount of government-supplied training under the New Deal and other youth programs. The 0.31 rate also falls well below the rate of 0.41 found for the

[35]The subsequent UK Household Longitudinal Study (UKHLS) does not contain any training questions in any of its waves.

U.K. by Arulampalam et al. (2004a) using the ECHP, despite that study's narrower definition of training. In the BHPS, respondents report that only 42% of the training spells lead to recognized qualifications, implying a rate of formal training receipt (using our definition from Section 2) of 0.13. The respondents also report that more than 84% of the training provides general skills, and that about 90% is either financed by the employer or takes place at the firm during working hours (with a notably lower percentage for training leading to a qualification). Conditional on participation, and consistent with the general pattern across studies, training duration has an extremely skewed distribution, with a mean of 20.6 days and a median of four days for "job-induction training", and a mean of 11.5 days and a median of two days for "current job skills" training. The authors do not find any economically important (unconditional) differences in training take-up, type of training, or financing between men and women.

4.3.2 Evidence from employer surveys

Bassanini et al. (2007) study training investments made by firms using the Continuing Vocational Training Survey (CVTS) from 2000.[36] The CVTS is a nationally representative employer-based survey targeting private-sector enterprises with at least 10 workers. Eurostat has carried out the survey every five years since 1995 in all EU member states plus Norway. The survey suffers from very low response rates in member states that leave it voluntary. For example, response rates in 2000 equalled 11% for Danish firms, 19% for Irish firms, 20% for UK firms, and 32% for German firms. Other member states, such as Romania and Lativa, achieve response rates above 90% by making such surveys mandatory.[37] The simple average of the national CVTS response rates equals 57%, only slightly lower than the 64% achieved in the EQW-NES (Bundesinstitut für Berufsbildung et al., 2005). The CVTS includes questions covering a wide range of firm training investments.

[36] The authors embed the analyses described here, along with an analysis of the ECHP similar in spirit to that in Arulampalam et al. (2004a), within their general survey of the European firm training literature.

[37] We do not know whether firms in these countries were also required to report the truth, the whole truth, and nothing but the truth, in mandatory surveys.

Among the training questions, the 2000 CVTS asked firms "During 1999 did any person employed by your enterprise undertake any of the following continuing vocational training activities?" where the first two of the seven options that are provided cover continuing vocational training courses managed either internally or externally.[38] Firms responding in the affirmative to either of those activities received follow-up questions asking about the total number of participants and paid working hours spent in either activity, along with fees and payments, travel and accommodation costs, labor costs for any internal trainers, and costs for room, equipment, materials and media.

Bassanini et al. (2007) report considerable dispersion in training expenditures measured as a percentage of total labor costs from as low as 0.5% and 0.8% in Romania and Poland, respectively, to as high as 2.8% in Sweden and the Netherlands and 3.0% in Denmark. Firms in the Nordic countries and the U.K. report that more than 50% of their workers participated in employer-sponsored training during 1999, whereas less than 10% of workers did so in Latvia, Lithuania, and Romania. Across all countries, smaller firms (10-50 employees) spend, on average, one percentage point less out of total wage costs on training than do large firms (at least 250 employees). As the authors note, fixed costs and scale economies in training along with easier internal reallocation of workers in large firms likely drive this difference.

4.3.3 Summing up

The European evidence sings in harmony with the North American evidence by highlighting the difficulties of interpretation and comparison that result from differences in question wording, in training type definitions, and in other dimensions of measurement across surveys and across studies. It further echoes the ubiquity of heterogeneity in training incidence and duration as a function of worker and firm characteristics. Taken

[38]The full list of options: a) Continuing vocational training courses managed internally; b) Continuing vocational training courses managed externally; c) Planned periods of training, instructions or practical experience, using the normal tools of work, either at the immediate place of work or in the work situation; d) Planned learning through job rotation, exchanges or secondments; e) Attendance at learning / quality circles; f) Self learning through open and distance learning; g) Instruction at conferences, workshops, lectures and seminars.

together, the North American and European literatures combine to motivate our analysis in the following section, which compares training across countries using a common data source.

4.4 New evidence from PIAAC

In this section we describe participation patterns across a host of countries using the publicly available micro data from the *Programme for the International Assessment of Adult Competencies* (PIAAC). PIAAC fielded a major survey in each of 38 countries between August 2011 and April 2018. The PIAAC survey aims to measure key skills such as literacy, numeracy, and problem solving in nationally representative samples. The survey also collects information on how adults use these skills at home and at work and, what matters for our purpose here, what kind of human capital investments they took part in during the 12 months prior to the interview. The response rate averages 62% across the participating countries, with Greece at 41% and Sweden at 45% at the low end and Turkey at 80% and Peru at 83% at the high end. The total achieved sample exceeds 250,000 respondents.

The OECD provides PIAAC public use files (PUFs) for 33 countries: Austria, Belgium (Flanders only), Chile, Czech Republic, Denmark, Ecuador, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Japan, Kazakhstan, South Korea, Lithuania, Mexico, the Netherlands, New Zealand, Norway, Peru, Poland, Russia, Singapore, Slovak Republic, Slovenia, Sweden, Turkey, United Kingdom (England and Northern Ireland), and the United States.[39]

For each country in our analysis, we select respondents ages 25 to 59, not self-employed, who report working full-time in a private or non-profit sector job at the time of the interview. We condition on employment only at the time of the interview, not over the entire 12-month reference period for the training question. As a result, we will pick up some non-firm training. Imposing these restrictions on the population of

[39] The PUFs from Canada and Spain lack the training questions central to our purpose. Australia, Cyprus, and Indonesia do not offer PUFs. Travel to oecd.org/skills/PIAAC/data for PUFs as well as documentation.

interest leaves us with a total sample of 47,226 potential trainees.[40] We impose our age limits to avoid the period of formal schooling at the start of the working life and and to avoid complications with cross-country differences in normative retirement behavior at the end of it. We focus on the training patterns of full-time, private-sector workers as non-employed workers, the self-employed, and government workers typically face very different training choice problems. Even after imposing these boundaries on our population of interest, our sample still contains more observations in more countries than in any of the other studies investigating the incidence, extent, and correlates of training.

Immediately after collecting information on the respondent's age, gender, and educational attainment, the PIAAC survey asks "During the last 12 months, that is since <insert month and year 12 months prior to the interview>, have you studied for any formal qualification, either full-time or part-time?" In the U.S. 13% respond in the affirmative, compared to 16% in the U.K. and just 3% in Germany. Germany's rate lies well below the OECD average rate of 9% and the EU average rate of 8%. The US rate from the PIAAC falls slightly below the 17% rate obtained by Loewenstein and Spletzer (1999a), though differences in the populations covered by the PIAAC and the NLSY-79 and in their definitions of formal training cloud the comparison. The UK rate resembles that found in Booth and Bryan (2007).[41]

[INSERT TABLE 3 HERE]

The survey then asks a set of detailed questions about the most recent formal training spell.[42] These include ISCED-based questions on the level of training, as well

[40]The sample loss journey from 250,254 to 47,226 includes 53,050 from countries without PUFs or without a training question, 59,902 due to our age restrictions, 55,434 due to our full-time work requirement, 22,135 public sector workers, and 12,507 self-employed workers.

[41]The PIAAC training question generally has a low item non-response rate. Only Belgium (0.052), the Netherlands (0.021), Norway (0.022), and the U.S. (0.040) have rates above 0.020.

[42]Considering only a single training spell apparently does not lose that much information as 81% of those who report participating in training during the reference period have only a single spell, and 10% have only two spells.

as questions about field of study, whether the training was job-related or not, who financed the training, employment status while taking the training, and whether the training took place during working hours or not.[43] Alas, the PIAAC survey does not collect information about hours of formal training (though it does collect information on hours of informal training, which we describe below). We use the responses to these questions to classify training spells based on the employment status of the respondent at the time, based on whether or not an employer helped finance the training, and based on whether or not the respondent viewed the training as job-related.[44]

Descriptive statistics on formal training from the PIAAC appear in the upper rows of Table 3. We present separate statistics for the U.S., the U.K., and Germany followed by the overall average of the OECD countries for which we have data in the right-most column. That average weights each country equally. The first row presents the overall rates noted above. The next two rows condition on employment status at the time of training. Not surprisingly given that sample inclusion requires employment at the time of the interview, the vast majority of the formal training occurs when respondents have a job. Germany exhibits the smallest difference, due almost entirely to a lower rate of formal training while employed.

The rate of reported employer financing of formal training varies widely across countries, as shown in Figure 1. The U.K., Denmark, Ireland, and the Netherlands lead the pack at 73%, 73%, 75%, and 79%, respectively. The average among our countries equals 57% (= 5.2 / (5.2 + 3.9) x 100) from Table 3. Respondents code the vast majority of formal training incidents as job-related. We do have some worries about how respondents think about the employer financing question as, e.g., in the United States more than half of all training financed at least in part by the employer gets reported as not job-related.[45]

[43]ISCED = International Standard Classification of Education, a project of UNESCO.

[44]The PIAAC survey, like many others, explicitly tells the respondents to think of training that replaces work activities as employer financed, along with training that takes place outside the respondent's regular working hours but where the trainee either receives payment for the extra hours or time off work. The survey also reminds respondents that "job-related" may refer to a specific job or to general employability.

[45]This could reflect poor targeting of training by some employers and/or some workers applying a narrow, task-oriented definition of job-related that omits things like team-building or diversity training.

[INSERT FIGURE 1 HERE. Caption: Formal training incidence across countries. Note: The formal training measure is taken from the PIAAC sample of 25-59-year-olds working full-time in the private sector. OECD averages are (unweighted) country averages of available OECD member states.]

[INSERT FIGURE 2 HERE. Caption: Informal training incidence across countries. Note: The informal training measure is taken from the PIAAC sample of 25-59-year-olds working full-time in the private sector. OECD averages are (unweighted) country averages of available OECD member states.]

Following its formal training questions, the PIAAC survey asks about four different types of what it labels *organized learning activities*, which we take the liberty of calling "informal training", as in combination they roughly correspond to our definition in Section 2. These types consist of "courses conducted through open or distance education", "organized sessions for on-the-job training or training by supervisors or co-workers", "seminars and workshops", and "courses or private lessons". Two in three Americans in our sample report having participated in at least one such informal training activity during the 12-month recall period, compared to the OECD average of 54%. Figure 2 presents the rates for all of the countries in our data. They differ considerably, from less than 20% in Kazakhstan and Russia, to around 30% in Italy and Greece, up to almost 70% or more in the Nordic countries, the Netherlands, and New Zealand. Nearly all of the informal training is job-related, occurs while the respondent is employed, and is financed at least in part by the employer. Organized on-the-job training has the highest frequency among the component activities, with around half of respondents in the U.S., U.K., and Germany reporting participation in the last 12 months. Indeed, one in three U.S. and U.K. respondents report participating in at least one workshop or seminar over this period.

Table 3 shows conditional hours of job-related informal training and other informal training. Adding them up, informal training averages 113 hours among those who receive it in the U.S. and 101 hours in the OECD overall. This implies that the average recipient of informal training experiences nearly three weeks of it in a year. Once again, average hours disguise a huge amount of heterogeneity in reported hours among respondents.

Finally, Figure 3 and the last two lines of Table 3 present statistics on the incidence of casual training and of learning-by-doing. The underlying survey questions (which appear in Figure 3) differ from the questions on formal and informal training already considered in two important ways. First, they do not offer the respondent a reference period. Second, the responses take the form of a Likert scale with the options “daily”, “every week”, “every month”, “less than once a month” and “never”. Table 3 shows the percentages of respondents who do not choose “never” as their response. The most striking result concerns the fraction of respondents in some countries who *do* claim to “never” have learned a skill from a co-worker or supervisor and/or “never” to have experienced learning-by-doing. For example, as many as one in five Kazakhstani, Russian, and South Korean workers claim never to have received casual training and never to have experienced learning-by-doing. We suspect question wording and/or translation issues play a role in these puzzling responses.

[INSERT FIGURE 3 HERE. Caption: Casual training and learning-by-doing across countries. Note: The casual training and learning-by-doing measures are taken from the PIAAC sample of 25-59-year-olds working full-time in the private sector. OECD averages are (unweighted) country averages of available OECD member states.]

The first two columns of Tables 4 document the mean characteristics of participants and non-participants in formal training in the US PIAAC data, while the third column presents a standardized difference. Table 5 does the same for informal training. We divide the characteristics into three sets: worker characteristics, job characteristics, and firm characteristics. The two last columns in each table contain coefficient estimates and associated p-values from linear probability models of training incidence that include all of the variables in Tables 4 and 5. In notation, we estimate versions of

$$Y_{us} = \beta_0 + \beta_{worker,us}X_{worker,us} + \beta_{job,us}X_{job,us} + \beta_{firm,us}X_{firm,us} + \epsilon_{us} \tag{6}$$

where Y_{us} denotes an indicator for one specific training type in the PIAAC data, and $X_{worker,us}$, $X_{job,us}$, and $X_{firm,us}$ denote the worker, job, and firm covariate sets.[46] We present p-values from tests of the null that the coefficients on particular variables (where a particular variable may enter as a set of indicators for categories) equal zero, as well as the nulls that all of the coefficients on variables in one of the three sets equal zero. The table notes also provide the partial R^2 values associated with each of the three sets of characteristics. We estimated related models that start with one set of characteristics and sequentially add the other two, but do not tabulate them for reasons of space. We do mention a couple of findings from these models.

We interpret the estimates in these models as descriptive rather than causal. Many of the conditioning variables, particularly some of those in the job characteristics set such as what skills the worker uses, whether they manage other workers or not, and their relative wage, surely arise jointly with the training outcomes. Instead, we view the multivariate results as helping to sort out which of the univariate differences result from one determinant of training serving as a proxy for another. We think that the overall sense that our estimates provide of the relative importance of worker, job and firm characteristics adds value, as does the analysis of cross-country heterogeneity in the determinants of training that we present later in the subsection. In a broad sense, we view our findings in this section as inputs into future work, both on the applied theory front and the causal empirical front.

[INSERT TABLE 4 HERE]

[46] We recode item non-response values of X to zero and include indicators for missing values. To save space, we omit the associated estimates in the tables.

[INSERT TABLE 5 HERE]

We now consider the patterns revealed in the data for each set of conditioning variables—worker characteristics, job characteristics, and establishment characteristics—in turn. Taken together the PIAAC variables produce R^2 values of 0.109 in the linear probability model for formal training and 0.245 for informal training. We expect that the difference in R^2 values arises in part from the fact that the unconditional rate of participation in informal training lies much closer to 0.50.

Unconditionally, women have much higher probabilities of participation in both formal and informal training than men. These large differences essentially disappear with conditioning. Our finding using recent data differs from the (much) earlier analyses for the U.S. in Lynch (1992) and BBB (1993), both of which find men conditionally more likely to participate. We think the difference represents a real and profound change in the role of women in the labor market relative to the labor markets that generated their data. Due to confidentiality concerns, the PIAAC survey does not include information about race in its public use files.

As predicted by the standard theories laid out in Section 5, the unconditional and conditional probabilities of formal training decline strongly with age. Relative to the reference group of 35-44 year-olds, respondents ages 25-34 have a 0.10 higher conditional probability of receiving formal training, and respondents ages 45-54 a 0.06 lower one. These differences prove highly statistically significant. Our respondents, on average, appear to agree that it makes sense to invest in skills when younger in order to bear a lower opportunity cost and to experience a longer payoff period. In marked contrast, we see much smaller differences by age for informal training, differences that fail to attain conventional levels of statistical significance.

Perhaps surprisingly, the presence of a spouse decreases the unconditional probability of formal training but modestly increases the unconditional probability of informal training, the opposite of what one would expect if a spouse relaxed credit constraints around investments in formal training. Both differences shrink substantially with conditioning (even conditioning only on the personal characteristics), and do not attain conventional levels of statistical significance. The presence of children does not matter much either. Being born in the United States increases the unconditional probabilities but moves the conditional probabilities only a little.

The educational attainment measures we consider draw on two questions from the PIAAC survey. The first asks for the highest level of completed education. We collapse the responses to this question into three categories—primary and lower secondary, upper secondary, and tertiary—and include indicators for two of them in our multivariate models. In the U.S. context these categories (typically) correspond to completing eighth grade, completing high school, and completing some post-secondary degree or certificate. The second question asks if the respondent ever began a degree or certificate program but did not complete it. We code this as an indicator equal to one for respondents who drop out. This combination requires care in interpretation in a multivariate context. For example, a respondent who completed a BA but dropped out of a doctoral program would have both the tertiary indicator and the dropout indicator equal to one. In the data, most (about 65 percent) of those coded as dropouts finish secondary school and then start but do not finish a post-secondary program.

The data reveal strong associations between schooling and training. Unconditionally, take-up of both formal and informal training strongly increases with educational attainment, while having dropped out of school at some point decreases the incidence of formal training and increases the incidence of informal training. These bivariate patterns resemble those found in earlier studies. Things change in the multivariate analyses. For formal training, the point estimates on the schooling attainment indicators become close to zero and lose their statistical significance as soon as we add the job-related characteristics. However, a substantively and statistically meaningful negative conditional relationship (-0.04) remains for the drop out indicator. For informal training, respondents with tertiary education have a much higher (0.10) conditional probability of training than high school completers (our omitted group). We also obtain a strong and statistically significant positive coefficient (0.07) on the dropout indicator. We do not have a story for the informal training results worth sharing with the reader. Our failure to find a conditional correlation between schooling attainment and formal training differs from findings in many of the earlier U.S. studies reviewed above. We expect this reflects some combination of changes in the labor market (and in the skill production function) and differences in conditioning sets.

Worker characteristics matter, especially for the take-up of formal training. The partial R^2 of the worker characteristics in the formal training model equals 0.045, or nearly half the overall R^2 of 0.109. Age and high school dropout status generate much of the explanatory action for formal training. In contrast, though the worker characteristic variables pass a collective F-test with a p-value of 0.04, they matter much less in a substantive sense for informal training, with a partial R^2 of 0.014 relative to an overall R^2 of 0.245. Educational attainment accounts for much of the explanatory power of the worker characteristics in the informal training regression.

A module of the PIAAC survey collects information on skill use at work for four separate sets of skills: information and communication technology (ICT, seven measures), numeracy (six measures), reading (eight measures), and writing (four measures). We sum and then standardize the measures for each skill set to have mean zero and standard deviation one. Other job characteristics provided in the PIAAC survey include an indicator for managerial responsibilities, occupation (four categories)[47], and indicators for the decile of the national wage distribution in which the individual falls.[48] At the margin, these job characteristics increase the R^2 of the linear probability model for formal training by 0.03 and the R^2 of the linear probability model for informal training by 0.80. These values indicate similar relative contributions given the much higher overall R^2 of the model for informal training. We easily reject the joint null of zero coefficients on all of the job characteristic variables.

[47]These draw on the International Standard Classification of Occupations (ISCO), a project of the International Labor Organization (ILO).

[48]The OECD constructs the wage decile indicators using information on hourly earnings (inclusive of bonuses) that it in turn constructs for the entire survey sample in each country based on underlying responses on the timing and amount of compensation.

Collectively, the skills used at work variables matter, as we can reject the joint null of zero coefficients on all of them at the 0.01 level for both formal and informal training. Having a job involving writing strongly predicts both formal and informal training incidence, while having a job involving reading only strongly predicts informal training. Oddly, having a job involving ICT has a large and statistically significant negative coefficient in the linear probability model for formal training.

Managing others has a negative correlation with formal training incidence both unconditionally and conditionally. The conditional relationship is relatively large at -0.04 and statistically significant. Of course, managers often become managers because they already have lots of skills, and so tend to serve as the trainer rather than the trainee. In contrast, managing others has a strong positive unconditional relationship with informal training that changes sign to a weak negative relationship with conditioning. Looking by occupation category, we find that although skilled workers have a higher unconditional probability of receiving both formal and informal training, this relationship does not survive the conditioning. Like others before us, we find lower training probabilities for blue-collar occupations, a pattern that persists with conditioning, particularly so for formal training with a substantial, and statistically meaningful, point estimate of -0.07 vis-à-vis the mean of 0.13.

We cannot reject the joint null hypothesis of equal coefficients on the wage decile indicators for either formal or informal training. For formal training we see quite small conditional and unconditional differences in incidence, other than a dip for the top decile. Given that respondents report that much formal training does not receive any employer support, this constitutes quite strong evidence against the importance of credit constraints for participation in formal training on the job. For informal training we observe substantively large conditional differences for the lowest decile and the top three deciles, with informal training incidence generally increasing in the wage level both conditionally and unconditionally.

The bottom sections of Tables 4 and 5 display our estimates for the limited set of establishment characteristics available from the PIAAC survey. Collectively, these variables matter in both the statistical sense, as we can reject the null hypothesis that all of their coefficients equal zero with p-values below 0.01 in both cases, and in the substantive sense that they contribute meaningfully to the overall explanatory power of the model, particularly in the case of formal training. Looking at particular establishment characteristics, size correlates with informal training provision but this pattern mostly disappears with the conditioning. Whether the establishment forms part of a larger firm strongly predicts receipt of informal training both unconditionally and conditionally. Finally, we see large differences in formal training participation across industries, with many of the coefficient estimates in the linear probability model exceeding (in absolute value) the unconditional mean of the dependent variable. We find less variation across industries in the incidence of informal training, perhaps because it is so common to begin with.

Finally, we investigate heterogeneity in the determinants of formal and informal training among the 33 countries with available PIAAC data. We do so by estimating

$$Y_c = \beta_0 + \beta_{worker,c} X_{worker,c} + \beta_{job,c} X_{job,c} + \beta_{firm,c} X_{firm,c} + \epsilon_c \tag{7}$$

where $c \in \{1, ..., 33\}$ indexes countries. Rather than overwhelming ourselves and the reader with 33 sets of estimates along the lines of those in Tables 4 and 5, we instead present summary statistics on the estimated coefficients from the 66 models (33 countries times two training types) in Table 6 (for formal training) and Tables 7 (for informal training). The variation across the countries in the data in educational institutions, in how school-to-work transitions happen, in labor market flexibility, and so on lead us to expect heterogeneous coefficients for at least some characteristics. For example, education may matter more in countries with strong tracking in their primary and secondary school systems.

The key econometric issue in this context arises from the fact that taking the variance of a set of estimates always over-estimates the variance of the underlying parameters because the variance of the estimates combines the variance of the underlying parameters with the sampling variation. Expressing a specific estimated coefficient for country c as a combination of the underlying population value and the estimation error gives

$$\hat{\beta}_c = \beta_c + v_c \tag{8}$$

where v_c denotes the estimation error. Then for the variance of the estimated coefficients on a particular characteristic we have the decomposition

$$var(\hat{\beta}_c) = var(\beta_c) + var(v_c) \tag{9}$$

Our country-specific coefficient estimates vary in their standard errors, both because the PIAAC sample sizes vary across countries and because the amount of residual variation (or, put differently, the explanatory power of our covariates in the linear probability model) vary across countries.

We adopt standard methods that "shrink" the estimated coefficient variances based on an implicit estimate of $var(v_c)$. Specifically, we implement the framework provided by Böhning et al. (2002), who write in the context of the meta-analysis literature.[49] Doing so implies that we think (not unreasonably) about our 33 estimated models for a particular training type as separate studies in the sense of that literature. Implicitly, the estimator shrinks precise estimates less than it shrinks imprecise estimates when constructing its consistent estimate of the variance of the underlying population parameters.

The first column of estimates in Tables 6 and 7 consists of the simple mean of the country-specific estimates of the coefficients on that row's characteristic. We take the simple average because the country represents the unit of analysis for this exercise. The second column presents the p-value from a test of the null that the underlying coefficient has the same value for all of the countries in the data (i.e., that the "common coefficient" model holds for this characteristic). We have a strong prior that this

[49] This same methodological issue comes up in the estimation of the standard deviation of teacher value-added in the primary and secondary education literature. Guarino et al. (2015) and Araujo et al. (2016) provide methodological discussion in that context.

null never literally holds in the data. As such, we encourage a casual interpretation of the p-value as a quantitative measure of the strength of the evidence against the null for a particular conditioning variable. The third column presents the estimated standard deviation of the underlying country-specific coefficients, i.e. $\widehat{sd(\beta_c)}$, following "shrinkage" to remove the variance component resulting from sampling variation in the individual country-specific estimates. Finally, the fourth column presents the ratio of the shrunken estimate of the variance to the unadjusted variance of the estimates (i.e., $\widehat{var(\beta_c)}/var(\hat{\beta}_c)$).[50] We report zeros in both the third and fourth columns when we obtain a non-positive estimate of the shrunken variance. While the population variance necessarily has a non-negative value, the estimated shrunken variance will sometimes take negative values in finite samples due to some combination of a small population variance, sampling variation, and failure of the assumptions underlying the shrinkage estimator. Quite reasonably, relatively high p-values on the test of the null of the common effect model in the second column always accompany the zeros in the third and fourth columns.

[INSERT TABLE 6 HERE]

[INSERT TABLE 7 HERE]

For reasons of space, our discussion of the results in Tables 6 and 7 explores the patterns on the worker characteristics in some detail but only provides highlights from the job and establishment characteristics. To start, we fail to detect any conditional difference in the take-up of either formal or informal training between men and women. The average point estimate lies very close to zero in both cases, and fails to achieve statistical significance. For formal training, we fail to reject the null of the common effect model, i.e. the null of no difference in the coefficients on the male indicator within

[50] In the notation of Böhning et al. (2002), the second column equals $\sqrt{\hat{\tau}^2}$, the third column provides the p-value associated with their Q-statistic, and the fourth column presents I^2.

our disparate set of 33 countries. For informal training, we reject the common effect null but obtain a substantively very modest (0.03) estimate of the standard deviation of the population coefficients.

We detect a very strong negative age profile for participation in formal training. The average conditional difference in the coefficient on the indicator for age 25-34 equals 0.053, and differs statistically from zero at the 0.01 level. Similarly, the average coefficients for the indicators for ages 44-54 and 55-59 turn out negative and statistically different from zero, where, as in Tables 4 and 5, the omitted category consists of individuals ages 35-44. We can easily reject the null of the common effect model for the coefficients on the indicators for the youngest and oldest age groups. In both cases, our estimate of the standard deviation of the underlying population coefficients turn out somewhat smaller than the average. If we assumed a normal distribution of population coefficients, and it is not clear that we should, almost all would have the same sign as the average. For informal training, we find a similar age gradient in the average coefficients, but much less evidence of heterogeneity.

Marriage predicts a lower average chance of participating in formal training. Though modest at -0.015 it attains conventional levels of statistical significance. We see little relationship between marriage and informal training, and little evidence of coefficient heterogeneity for either type of training. Foreign-born (i.e. non-native) respondents have a lower conditional probability of informal training, though we cannot statistically distinguish the average of -0.022 from zero. We do find strong evidence that the coefficients on the native indicator differ across country for informal training, which makes sense given large cross-country differences in the size and features of the foreign-born population. We see no relationship between the native indicator and formal training, and no evidence of heterogeneity either, though we expected to see it here too.

Not having any children increases the conditional probability substantially for both formal (0.029) and informal (0.021) training; both coefficients differ statistically from zero, though only for formal training do we find evidence of heterogeneous coefficients. For informal training, but not for formal training, the presence of at least one child under six years of age further reduces (-0.021) the probability. As those with lived experience of children will know well, they impose serious time and flexibility constraints. To the extent that either type of training takes place at odd hours or at a location other than the workplace, those constraints may bind. When we estimate our linear probability models separately by sex (results not shown), we find both childless men and childless women have higher probabilities of training participation, but the point estimate for women equals nearly twice that for men for formal training. Having at least one child under the age of six has no association with the father's training participation but lowers training incidence considerably for women, by about three percentage points for formal training and by about five percentage points for informal training. Somewhat to our surprise, given notable differences among our countries in gender role norms, female labor force participation, and so on, these relationships do not appear to vary much across countries, as we never come close to rejecting the common coefficient null.

Also to our surprise, the cross-country average coefficients for the education variables imply some of the same qualitative patterns as in the US: a strong, positive education gradient in informal training, a weaker but still positive education gradient in formal training, a strong negative association between dropout (at any level) and formal training, and a strong negative relationship between dropout (at any level) and informal training. At the same time, the data strongly signal that the related coefficients vary across countries, often quite a lot. The variability does *not* surprise us at all given the rich assortment of educational and labor market institutions present in the 33 countries we study.

Among the job characteristics, the patterns associated with the wage deciles strike us as most interesting. For formal training, the coefficient averages qualitatively repeat what we found for the U.S. alone in Table 4, namely that rates of formal training incidence differ little over the wage distribution and, to the extent that they do, we tend to see more formal training at the low end of the distribution. Moreover, the data suggest little heterogeneity in these relationships across the countries in our data. This constitutes strong evidence against the empirical importance of credit constraints in worker and firm choices regarding formal training in developed countries. On the other hand, we find a strong conditional association between wage rates and the incidence of informal training. This relationship also appears relatively homogeneous across countries. As firms tend to pay for informal training, we interpret this as strong evidence of the ubiquity of positive selection into informal training.

Finally, among the establishment characteristics, we highlight the findings for the indicator for whether or not the establishment belongs to a larger firm and for the industry indicators. Like other studies before this, we find a clear positive relationship between being part of a larger firm and the incidence of both formal and informal training. We find no evidence of heterogeneity in this relationship across countries. We also find small, substantively meaningful, and statistically significant differences by industry in conditional rates of both formal and informal training. The evidence implies that the industry differentials differ across the countries in our data.

4.5 The big picture

We conclude Section 4 with a brief compendium of key substantive takeaways and "known unknowns" that signal opportunities for new theoretical and empirical research.

Asked to venture a definition of what applied labor economists mean when they talk about firm training, an economist unfamiliar with the literature just reviewed might describe something quite distinct from formal schooling that takes place solely within firms. In fact, quite a lot of firm training resembles formal schooling either in the sense that it comes with certificates or formal qualifications or even two-year college degrees attached to it or in the sense that it transpires outside the physical confines of the firm at an educational institution or training center.

The dimensions of training studied in the literature on the incidence and determinants of training just reviewed (which roughly correspond to the typology in Section 2) differ dramatically from those emphasized by the theoretical literature (and related empirical studies) that we engage in Section 5. That literature, dating back at least to Becker (1962), centers on the specificity of training, i.e. on the extent to which the skills created by the training do or do not have productive value at other firms. Indeed, as we will show, it focuses (until recently) on the stark binary distinction between fully general training, of value at every firm, and fully specific training, of value only at the training firm. While formality and generality overlap to some extent as formality correlates with verifiability (e.g. via recognized qualifications) and thus with transferability, the substantive distinction between the two literatures remains notable.

Whether employing a typology along the lines of that in Section 2, or adopting a binary split that distinguishes general and specific training, the existing literature pretty much always works with highly aggregated training types. We could learn much more about the nature, extent, and determinants of firm training by distinguishing more finely along multiple dimensions based on content, formality, provider, generality, duration, and so on. Alas, such developments will require data superior to that to which researchers currently have access. We have more to say along these lines in Section 5.8.6, which summarizes the key takeaways from the "training as treatment" literature.

Firm training is ubiquitous in the US and in other OECD countries as well. Over the course of a year, more than one in ten private sector workers in the US participate in some kind of formal training, two in three participate in informal training, and more than half report having participated in organized on-the-job training. The evidence from the literature and our own analysis of the PIAAC data effectively dispels the moral panic around an imagined lack of firm training for US workers relative to those in other industrialized countries. Indeed, American workers rank at or near the top of the PIAAC league table in terms of both formal and informal training, though measurement concerns (as exemplified by the low rates of measured firm training in Germany and Japan) restrain us from pushing this line too hard. The lack of comparable cross-country data on training in earlier periods prevents us from making claims about the persistence of the strong relative US performance.

Demographics matter when it comes to formal training. Around the OECD, younger workers receive relatively more formal training, as do (in most studies) workers who enter the labor market with a stronger educational preparation. In surprising contrast, demographics do not predict informal training very well, especially in the US. The literature reveals little evidence of the importance of credit constraints in reducing access to formal training, though further research with better data, particularly better measures of credit constraints, could easily change this picture.

Firm training takes up a non-trivial amount of participants' time. For example, according to our estimates based on the PIAAC data, a US worker spends (on average) two (work) weeks per year in informal training alone (as about two-thirds participate in informal training for an average of 113 hours). The broader literature reviewed above corroborates this point, and also emphasizes the extreme cross-sectional heterogeneity in the amount of time spent on training, even conditional on training type. The sometimes large time costs of training help explain why we find evidence that time constraints limit the amount of formal training some workers undertake. This finding points to the value of research on the temporal interaction of investments in human capital via firm training with major demographic events such as marriage, child-bearing, and migration.

The literature, as well as our analysis of the PIAAC data, reveal important heterogeneity in the incidence and extent of firm training across industries for both formal and informal training in most developed countries. One obvious explanation is that some industries find themselves better served by the regular schooling and college systems than others and hence need to make fewer post-schooling investments in industry-specific skills. Another obvious explanation notes that certain industries experience more rapid technological change than others and hence have a greater need for continuously equipping their workforce with new skills. We see plenty of room for research that seeks to account for industry differences in firm training.

We know very little about temporal dynamics of training over workers' lives. Do some workers participate year in and year out while others never or almost never do, or do training events spread out over time such that all workers eventually "take their turn"? We see a parallel to the literature that worries about the temporal persistence in earnings shocks, wherein more persistence implies more inequality in discounted lifetime earnings. In the firm training context, the greater the extent to which past training predicts future training, the more inequality in lifetime human capital accumulation. Decomposing the serial correlation in firm training into components at the individual, occupational, and industrial levels would also add value, as would studying the temporal relationship between receiving firm training while employed and receiving government training while not employed.

Finally, we know very little about the amount of resources devoted to training workers at firms. This is true in the narrow sense that we lack systematic data on the direct (e.g. payments to contract providers of training) and indirect (e.g. time workers spend training while "on the clock") costs incurred by individual firms. It is also true in the broader sense that we lack good estimates of the total amount of societal resources devoted to the production of human capital via firm training, estimates that we might compare with existing estimates of the amount of resources spent on human capital development via primary, secondary, and tertiary schooling or via investments by parents and the government prior to formal school entry.

5 Theories and evidence

5.1 Beckerian theory

5.1.1 Is training a choice?

The theory of human capital has a long history in economics, dating back at least to Smith's (1776) *Wealth of Nations*. While Pigou apparently coined the term "human capital" in 1912, it was present in all but name in Smith's work. Spengler (1977) quotes Smith:

> "... the acquired and useful abilities of all the inhabitants or members of the society. The acquisition of such talents, by the maintenance of the acquirer during his education, study, or apprenticeship, always costs a real

> expence, which is a capital fixed and realized, as it were, in his person. These talents, as they made a part of his fortune, so do they likewise of that of the society to which he belongs. The improved dexterity of a workman may be considered in the same light as a machine or instrument of trade."

In what follows, we will describe the major theories and the associated empirical findings from the literature on firm-provided training. In these papers, the authors generally assume that training represents a choice variable for the firm and, at least implicitly, for the worker. But how free are firms (and their workers) to set the level of training?

Consider a worker hired as a cook at the proverbial greasy spoon diner, not a job generally consider "high skilled". Suppose our diner cares about training costs and would like to keep them low. Could it reasonably save on training costs by hiring a worker with no experience as a short-order cook, sending them to the kitchen without any training, and instructing them to learn by doing? Undoubtedly not. The cook needs to know, at some level, how to prepare the dishes on the menu. Given the choice to hire a worker without experience, the manager of the diner really has no choice (at least at the extensive margin) regarding training. Alternatively, the firm can economize on training costs by hiring an experienced cook. Depending on the nature of their previous experience, doing so could eliminate some or all of the diner's initial training costs. Of course, a trained cook will expect a higher wage than an untrained one, and will likely imply higher search costs as well.[51]

Stepping back for a broader view, in many workplace contexts, employers face not a one-dimensional choice of how much to train but rather a choice over bundles of worker experience, search costs, and training expenses, all of them uncertain ex ante to at least some degree. This is the fundamental insight of John Barron and his various co-authors in a series of papers in the 1980s and 1990s.[52] The firm's training decision is

[51] Adding yet another dimension of choice, the firm could change its menu to limit the number or complexity of the dishes its new cook needs to learn. We see this as both a change in the production technology and a move in product space, one that might not please the diner's regular customers!

[52] See Barron et al. (1985), Barron and Bishop (1985), Barron et al. (1987b), Barron et al. (1987a, 1989),

inextricably tied to their decisions about search behavior.[53] Stepping back even further, production technology choices, presumably made at longer intervals, strongly affect the context within which firms make their training and search choices. We encourage the reader to keep the firm's trade-off between hiring less experienced workers and training them more or hiring more experienced workers and training them less in mind as they read through our description of the literature. We think this represents one of many directions in the firm training-literature that cry out for further research.

5.1.2 Becker and early theoretical work

Despite the reverence that economists hold for Smith, it was not until the late 1950s and early 1960s that they began using human capital to model workers' earnings, starting with the publication of Mincer (1958), Schultz (1961), and Becker (1962).[54] In his 1964 classic *Human Capital*, Becker enshrined the distinction between general training—which had value in the marketplace—and specific training—which had value only to the incumbent firm. Of course, Becker and the other early writers understood these as illustrative extremes, much like perfect competition and monopoly in product markets. Important early contributions to theory include Oi (1962), Parsons (1972), Rosen (1972a,b), Mortensen (1978), Hashimoto and Yu (1980), Hashimoto (1981), Carmichael (1983a,b, and 1985), and Ohashi (1983). In these early years, major contributions were largely theoretical due to a lack of available data on firm-provided training.

In Section 3, we described the difficulties that continue to plague training measurement efforts, especially what we call casual training. We are by no means the first to highlight these challenges. As Brown (1989a) argues:

and Barron et al. (1997b, 1999).

[53]A parallel point applies to workers, who may direct their search toward jobs they can already do or toward jobs that will provide them with additional skills.

[54]Somewhat confusingly, the literature credits the "Chicago School" with the development of human capital theory. While Schultz was a faculty member at Chicago and Becker and Mincer were both students at Chicago (although Mincer subsequently completed his doctorate at Columbia), Mincer published his 1958 paper while on the faculty at the City College of New York (moving to Columbia the following year) and Becker's 1962 paper and 1964 book appeared while he taught at Columbia. He moved to Chicago in 1970.

> Obtaining information on the extent of training of the workforce is complicated both by conceptual problems and by difficulty in actually measuring those aspects of training which seem relatively well-defined. Much of the conceptual difficulty of measuring employer-provided training is due to the fact that an important part of such training occurs informally, on the job. While there are difficulties in measuring formal training, what we would like to measure is relatively well-defined . . . Informal training is produced jointly with the primary output of the worker, and is therefore more elusive. Workers learn from watching other workers, may share easier ways to do the work either while working or during breaks, and are indirectly "instructed" whenever a supervisor constructively criticizes their work. Knowing whether informal training is happening in any given week may be difficult to determine; one hopes that for most workers it never ends.

The extreme version of Brown's never-ending cascade of informal training corresponds to Rosen's (1972a,b) learning-by-doing model wherein expertise increases monotonically with experience, albeit subject to diminishing marginal returns.

While we do not want to strain our readers' patience even further by reviewing the immense literature on learning-by-doing, Rosen's papers do raise interesting issues as to what precisely constitutes on-the-job training. When workers gain expertise from simply doing their job, how should we measure on-the-job training? Should we ignore the explicit training programs that we labeled formal and informal training in Section 2, and even ignore what we called casual training, whether by managers and co-workers or via quality time with a training video, and instead simply concentrate on measuring experience on the job? Probably not, as nothing prevents formal training *and* informal training *and* casual training *and* learning-by-doing from all making important contributions to overall human capital accumulation. Indeed, the literature reviewed in this chapter makes it clear that all four types of on-the-job skill development matter.

5.1.3 General training

Many of these early studies, such as Becker (1962), Ben-Porath (1967), Mincer (1958, 1962, 1974), and Oi (1962), used multi-period dynamic models, but a simple two-period model will suffice for our purposes. We make four important simplifying assumptions:

(A-1) Labor markets are perfectly competitive;

(A-2) There is perfect information in the labor market;

(A-3) Labor is homogeneous;

(A-4) Capital markets are perfect in the sense that workers may save and borrow at a fixed interest rate.

Assumption (A-4) allows us to use Fisher's (1930) separation theorem so that workers simply seek to maximize their wealth. We also assume, for now, only general training, meaning that the skills training provides have equal value in any firm. This makes training in our initial model very similar to schooling in simple models of homogeneous schooling, except that workers acquire it on the job instead of in a school. While extremely simple mathematically, this model provides deep economic insights that the literature sometimes neglects. Moreover, despite its simplicity, it can aid our thinking in regard to the importance and specific consequences of potential market failures.

In the first period of our two-period model, the value of a worker's marginal product equals V_1 and they receive a wage w_1. The also receive training that costs $c(h)$ where h denotes a specific level of training. Defining h in this way implies a binary choice problem at the extensive margin of training. In the second period of our model, the value of a worker's marginal product equals V_2, with $V_2 > V_1$, and they receive a wage w_2. Workers share a common discount term, δ.

Our simple structural model yields a simple solution. As all training is general and the labor market is competitive, in the second period workers will earn the value of their marginal product, so that $w_2 = V_2$. Because the firm is indifferent between retaining workers it trained and hiring workers trained at another firm in the second period, the first period wage depends only on the productivity and training costs in period one. Thus, we have

$$w_1 = V_1 - c(h) \tag{10}$$

which implies that workers pay the full cost of training and reap the full return.

Will a worker take this contract? To make this determination, we need to specify the worker's alternative. Let V_a denote the constant marginal product of workers without training. The firm's zero profit condition requires them to pay $w_{1,a} = w_{2,a} = V_a$, so workers will accept the training contract if

$$w_1 + \delta w_2 \geq w_{1,a} + \delta w_{2,a} \tag{11}$$

with equality holding if we see some workers accepting the training contract and other workers accepting the no training contracts, an assumption we will maintain for a bit.[55] In period two, trained workers earn more than their untrained counterparts, $w_2 > w_{2,a}$, but they pay for that privilege by accepting a lower first period wage, $w_1 < w_{1,a}$.

Importantly, the information in the wage profiles suffices to recover both the cost of training, $c(h) = V_1 - V_a = w_1 - w_{1,a}$, and the returns to training, $V_2 - V_a = w_2 - w_{2,a}$. This matters because we generally think it easier to measure wages than to measure Brown's "elusive" training. Indeed, as we shall see, much of the early empirical work uses observed wage profiles to infer the training that has taken place. In our simple model, training functions just like schooling in the sense that workers pay the full costs (putting aside the heavy public subsidies to formal schooling) and reap the full benefits. If a worker changes firms, the training is perfectly portable. Indeed, the curious reader may wonder, "Why is this training not provided at schools?" We will return to this question when we discuss post-Beckerian theories.

Indeed, some models do not distinguish between investments in human capital that

[55] As written, our model predicts that either all workers prefer the job with training, no workers prefer the job with training, or all workers are indifferent between the two. Adding a random preference for training with median zero completes the model by specifying what happens when equality holds in (11) for all workers and would imply that half get trained and half do not.

take place at educational institutions and at firms. For instance, Ben-Porath (1967) provides a canonical model of investment over the lifecycle, but in his paper workers produce human capital by taking time away from productive activities. He posits a Cobb-Douglas production function whereby workers combine purchased inputs and their own time to produce human capital. He embeds this production function in a continuous time lifecycle model with perfect capital markets. In his model, the worker's stock of human capital varies over time due to both investment via the production function and exogenous depreciation. As this description suggests, the model has the "look and feel" of models of investment in physical capital, with the important difference that the Ben-Porath model has a fixed time horizon T, reflecting our sadly finite lifetimes, rather than the infinite time horizon commonly assumed in models of investment in physical capital.

As Ben-Porath (1967, 364) notes, "The reader probably does not have to be told how simplifying many of the assumptions are and how many additional aspects of investment in human capital and of the determination of the life cycle of earnings are relevant." Yet, despite its simplicity, the Ben-Porath model produces tremendous insights. For instance, Ben-Porath demonstrates that the optimal investment path involves a period at the beginning of the life-cycle wherein the worker devotes all of their time to human capital investment and so has zero earnings (which they can do thanks to the assumed perfect capital markets). While we might naturally label this period "schooling," Ben-Porath (1967, 358) does not apply that label, but does note that "the early period when no earnings are realized is certainly not inconsistent with the real world." After this period of specialization, workers split their time between working and further investments in human capital and so have positive earnings. As they age workers optimally spend more and more of their time earning and less and less of their time investing in human capital, implying an age-earnings profile that increases at a decreasing rate over most of the lifecycle. As workers approach the end of their careers (i.e., approach T), earnings begin to decline as optimal investment goes to zero and net depreciation kicks in.

What a remarkable paper! A simple wealth maximization model with a Cobb-Douglas production function for human capital generates predictions broadly consistent with the patterns observed in modern economies. While we can certainly worry about many of the assumptions, the model helps us understand basic patterns of human capital investment and lifecycle earnings. Modestly, Ben-Porath (1967, 364) argues that "the main purpose of this paper was to raise some more questions rather than provide definite answers," but students of human capital should read this paper and digest its insights.

Jacob Mincer's (1974) book *Schooling, Experience, and Earnings* is another classic in the early literature on human capital. Heckman et al. (2003) call the model of earnings that Mincer develops in the book "a cornerstone of empirical economics." Mincer constructed models amenable to empirical estimation on the limited computational devices available at the time, devices orders of magnitude less powerful than the phones we carry in our pockets today. The book represents the culmination of more than 15 years of Mincer's research into human capital, stretching from his dissertation through his years of collaborating with Gary Becker during their time together at Columbia.

In the theoretical section of his book, Mincer develops models of investments in schooling and of post-schooling investments in human capital. In the subsequent empirical sections, he examines both the returns to schooling and the returns to labor market experience, which he interprets as a proxy for on-the-job training. Mincer understood the difficulties associated with measuring on-the-job training, noting in Mincer (1974, 129) that in the context of his model, the "magnitude of the cumulated investment cannot be observed, but it is a concave function of experience. Hence, to expand the schooling model into a more complete earnings function, the linear schooling term must be augmented by a nonlinear, concave, years-of-experience term."

In the process of analyzing the data, Mincer develops his justly famous "Mincer earnings function" (or just "Mincer equation") that relates the logarithm of a worker's earnings to their schooling level and a quadratic in their labor market experience.[56]

[56] In data sets lacking measures of actual experience, researchers adopt potential experience, typically

The Mincer equation flows directly from his model of formal schooling and on-the-job human capital accumulation. Indeed, the Mincer equation represents one of the few cases in labor economics where a specific estimating equation follows directly from the model of optimizing behavior assumed to generate the data.[57]

5.1.4 Specific training

For theorists, general training lacks excitement, in part because it poses no interesting contracting issues. In contrast, as Becker (1962, 1964) points out, things become more complicated, and thus more interesting, when considering firm-specific training. Becker emphasized the empirical importance of firm-specific capital, which increases worker productivity at the training firm but not at any other firm. In notation, let $V_{2,a}$ denote a worker's second period productivity at non-incumbent firms while V_2 remains their productivity at their incumbent firm and $V_2 > V_{2,a}$, reflecting the presence of firm-specific human capital.

Suppose a firm offers the worker the same contract we described in the general training context, in which the worker pays the full cost and reaps the full benefit of training. In the specific training context, this contract subjects the worker to a potential hold-up problem, because they cannot benefit from their specific skills at other firms. To see this, note that if, at the end of the first period, a devious firm reneges on the initial promise that $w_2 = V_2$ and instead offers a wage, $w_{2,d}$ such that

$$V_{2,a} < w_{2,d} < V_2 \tag{12}$$

the worker faces a quandary: Should they quit? If they stay, the firm pays them more

defined at age minus schooling minus six, where six represents the age of school entry, as a proxy. This practice never fit well for women and has become less appropriate for men as their mid-life labor force participation rates have fallen.

[57] Heckman et al. (2003) undertake a detailed study of the Mincer model's empirical performance. While they find impressive support for the model in data from earlier (1940, 1950, 1960) censuses, the model misses important features of the data in later (1970, 1980, 1990) censuses. They also note that researchers seeking to estimate the causal effect of schooling on earnings, rather than seeking to estimate the parameters of the equilibrium from the Mincer model, should not include experience on the right-hand side of their estimating equation because schooling affects labor market participation.

than their best alternative but less than the promised V_2. If they leave, they receive only $V_{2,a}$, which is even less. On the other hand, if the firm offered to pay the full cost and reap the full return of specific training, the firm would then face essentially the same potential hold-up problem from their workers.

Becker's solution is for the firms and workers to share the costs and the returns to training. When they share in the costs and benefits, they both want to extend the employment relationship because they both obtain "quasi-rents" from it. Indeed, shared financing of specific training implies more stable employment relationships than a world with a fully competitive labor market and only general training, in which even small, temporary productivity shocks can generate separations. In the terminology of Walter Oi (1962), the specificity of training makes labor a "quasi-fixed" factor of production.

Becker recognized these issues. Indeed, we quote from Becker (1962, 20) in some detail because he expresses the underlying costs and benefits of specific training so clearly:

> "Firms paying for specific training might take account of turnover merely by obtaining a sufficiently large return from those remaining to counterbalance the loss from those leaving. (The return on "successes"—those remaining—would, of course, overestimate the average return on all training expenditures.) Firms could do even better, however, by recognizing that the likelihood of a quit is not fixed but depends on wages. Instead of merely recouping on successes what is lost on failures, they might reduce the likelihood of failure itself by offering higher wages after training than could be received elsewhere. In effect, they would offer employees some of the return from training. Matters would be improved in some respects but worsened in others, for the higher wage would make the supply of trainees greater than the demand, and rationing would be required. The final step would be to shift some training costs as well as returns to employees, thereby bringing supply more in line with demand. When the final step is completed firms

> no longer pay all training costs nor do they collect all the return but they share both with employees. The shares of each depend on the relation between quit rates and wages, layoff rates and profits, and on other factors not discussed here, such as the cost of funds, attitudes toward risk, and the desire for liquidity."

Becker clearly had in mind a model a bit more complicated than ours, and considerably more complicated than the formal model he actually wrote down. For instance, Becker's formal model has no turnover. It also leaves the "sharing rule" that specifies the division of the costs and benefits of training between the worker and the firm indeterminate, just as we do in equation (12). Much of the early literature sought to formalize the optimal sharing rule in contexts with investment in firm-specific training.

In order to make this a more interesting problem, we relax assumption (A-2) by introducing some asymmetric information. Toward that end, suppose that productivity at the incumbent firm for the i-th worker equals $V_2 + \epsilon_i$, where ϵ_i is the private information of the firm. It could include things like how well worker i gets along with their co-workers. Further assume that $\mathrm{E}(\epsilon_i) = 0$ and that firms learn ϵ_i at the end of the first period. With this one-sided asymmetric information, we can give the second-period wage setting power to the firm and still achieve the efficient outcome. The firm will simply set $w_2 = V_{2,a}$ and expropriate all the rents in the second period. However, the competitive labor markets assumed in (A-1) mean that the firm will pay for this privilege because the first-period wage will adjust to eliminate all of the expected economic profits from hiring the workers.

The contracting environment becomes much more complicated (and so of greater interest to theorists) when workers also possess asymmetric information. One simple scheme assumes that workers have idiosyncratic tastes for jobs that only they observe. In notation, let α_i denote the value of the i-th worker's idiosyncratic taste for some specific job. The i-th worker's utility at that job then becomes

$$u_{2,i} = w_{2,i} + \alpha_i. \tag{13}$$

We now have a model with two-sided asymmetric information. We have added an i subscript to the second period wage because, as we show in what follows, it will now vary across workers in some cases.

Hashimoto (1981) and Carmichael (1983a, 1985) develop models in this vein. Mortensen (1978) and Hashimoto and Yu (1980) introduce the asymmetric information by having workers' alternative employment opportunities subject to random shocks. Papers such as Parsons (1972) and Ohashi (1983) simply posit quit functions, but one could rationalize their quit functions using either the idiosyncratic tastes of workers for positions or random fluctuations in outside offers. Models of this type capture essential features of the employment relationship. Firms are uncertain whether their workers will continue to work for them, rendering investments in their workers inherently risky. Similarly, workers are uncertain whether their current firm will retain them or even continue to exist, so that investments in specific training embody inherent risks for them as well.

With this framework, we have a bilateral monopoly with asymmetric information, a classic problem in mechanism design. If the firm makes the wage offers, it will always lower the second period wage below the efficient level. While the lower wage raises profits, it also implies that some efficient exchanges do not take place as workers leave their incumbent firms for alternative firms despite lower productivity there. The problem persists if the workers make the wage offers, as they will offer wages above the level required for retention at the incumbent firm as doing so maximizes their expected utility. Either way, despite their individual optimality, these rent-seeking behaviors reduce the joint gains from the employment relationship.

Researchers suggested many clever solutions to this asymmetric information problem, including sharing rules as in Hashimoto (1981), separation penalties as in Mortensen (1979), and promotion ladders as in Carmichael (1983a). Alas, none of these ingenious ideas fully solve the problem. Indeed, Myerson and Satterthwaite (1983) show that, in the context of the model discussed here, *no* mechanism exists that will always achieve the efficient outcome. The asymmetric information necessarily entails inefficiencies.

Black and Loewenstein (1997) provide a thorough discussion of this important result.

At this point, a bit of formality will add value. Begin by again denoting productivity and wages in period one by V_1 and w_1, respectively. Equation (13) shows the worker's utility at their incumbent firm, with α_i still their private information. Workers remain at the incumbent firm in the second period when

$$w_{2,i} + \alpha_i \geq V_a \tag{14}$$

where V_a again denotes the worker's wage and productivity at the relevant alternative firm. In the second period, a worker's productivity at their incumbent firm equals $V_2 + \epsilon_i$, where ϵ_i remains private information for the firm.

We now face the issue of the contracting assumptions. We begin by assuming that firms make take-it-or-leave-it offers to workers, but we will talk about other contracting assumptions as well. If workers remain at their incumbent firms, firms earn profits $(V_2 + \epsilon_i - w_{2,i})$, but of course profits accrue only if workers accept their offers. The firms' problem then is to pick the second period wage offer, $w_{2,i}$, that maximizes expected profits, or

$$\pi_2 = \Big(1 - F(V_a - w_{2,i})\Big)\Big(V_2 + \epsilon_i - w_{2,i}\Big) \tag{15}$$

where $F(\cdot)$ denotes the distribution of workers' idiosyncratic tastes for their incumbent firms. A bit of manipulation yields the necessary condition for expected profit maximization,

$$V_2 + \epsilon_i - w_{2,i} - M(V_a - w_{2,i}) = 0, \tag{16}$$

where $M(\alpha) = \big(1 - F(\alpha)\big)\big(f(\alpha)\big)^{-1} > 0$ and $f(\alpha)$ is the density corresponding to $F(\alpha)$. Efficiency requires $w_{2,i} = V_2 + \epsilon_i$, so the addition of the $M(\cdot)$ term in equation (16) means that the firm sets the wage, $w_{2,i}$ too low for efficient turnover. This term arises because firms understand that the idiosyncratic tastes of workers for their incumbent firms represent rents, part of which they try to expropriate by offering a sub-optimal wage. This sub-optimal wage in turn implies that some jointly efficient production will not take place. In choosing its second-period wage offer, the firm balances this lost production against the expected gain from the expropriation of (some of) the worker's rents. The expropriation itself is just a rent transfer from workers to firms that has no impact on efficiency.

Given the firms' rapacious behavior, one might be tempted to let workers make the take-it-or-leave-it offers. Alas, workers, acting optimally, would engage in similar rent-seeking behavior by offering higher-than-optimal wages in order to capture some of the rents implicit in their firm-specific productivity. Because their idiosyncratic productivities are the private information of the firm, this rent-seeking behavior results in inefficient turnover.[58]

We now finish our discussion of the contract with firms setting wages by specifying how they set wages in the first period. The firms in our model recognize that they can expect profits of $\mathrm{E}_\epsilon(\pi_2)$ in the second period. In a perfectly competitive labor market, this implies that first period wages get bid up until

$$w_1 = V_1 - c(h) + \delta \ \mathrm{E}_\epsilon(\pi_2), \tag{17}$$

where $c(h)$ again denotes the training cost, as in equation (10).

In short, firms compete away their expected second-period profits, $\mathrm{E}_\epsilon(\pi_2)$, via their wage offers in the first period. If we could magically stop firms from engaging in their rent-seeking behavior, the starting wage would increase because $\mathrm{E}_\epsilon(\pi_2)$ would increase. Finally, we need to make sure that this wage contract is acceptable to the worker. The contract must at least offer the worker the same utility as their next best alternative, call it U_0, or

$$w_1 + \delta\Big(\mathrm{E}((1 - F(V_a - w_{2,i})(w_{2,i} + \alpha_i)|\alpha > V_a - w_{2,i})) + \mathrm{E}\big(F(v_2 - w_{2,i})V_a\big)\Big) \ \geq U_0 \tag{18}$$

[58]We can achieve an efficient contract with only one-sided asymmetric information. To see why, suppose firms know workers' α_i's. Then we can have firms set a wage $w_{2,i} + \alpha_i = V_a$, which implies only efficient turnover. Similarly, if workers' know their ϵ_i's then we simply let workers set wages to equal to $w_{2,i} = V_{2,i} + \epsilon_i$ if they wish to remain with their incumbent firm in the second period.

In period one, the worker has not yet learned α_i and the firm has not yet learned ϵ_i, so the constraint involves the worker's expected utility in period two.

Why does the inefficiency that results from two-sided asymmetric information matter? Anytime the market fails to achieve an efficient outcome, it makes certain investments infeasible. Without these investments, certain jobs will not be feasible under two-sided asymmetric information that would be feasible with fully-informed agents. If we look carefully at equation (18), anything that lowers the left-hand side makes it harder to meet the constraint. In this model with homogeneous workers, firms must reward workers for undertaking risky investments in specific training, where the risk arises from the possibility of turnover. Thus, anything that limits the sub-optimal turnover lowers the cost of training, which in turn results in more training opportunities for workers.

In the real world, unlike the world of our simple models, the sharp distinction between general training and specific training tends to soften. How many firms must demand a specific type of training before we label it general rather than specific? Should we have in mind a continuous notion of specificity rather than binary one? Perhaps industry-specific or occupation-specific human capital matters relative to firm-specific training (noting that one must then define the breadth of the relevant occupations and industries). In an important paper Neal (1995) demonstrates that displaced workers who leave their former industry earn less than those who find another job in the same industry. But as Neal (1995, 669-670) notes, "I must acknowledge the possibility that the results outlined above reflect the importance of skills that are not truly specific to given industries, but rather specific to a set of jobs that are associated with the intersection of certain occupations and industries." Similarly, Becker (1962) writes that "Much on-the-job training is neither completely specific nor completely general but increases productivity more in firms providing it and falls within the definition of specific training."

Indeed, the fact that the training occurs at the firm (or at least via the firm) rather than at a school should give one pause regarding claims of generality. High schools, community colleges, and universities all offer a wide variety of vocational training today (much to the horror of some of our colleagues), and one may wonder why firms, which clearly have not specialized in the production of human capital, would choose to provide general training rather than leave it to educational institutions that have specialized in it. We can think of a couple of reasons. First, the firm may think it can provide higher quality training at the same or lower cost. Relying on a school may imply contracting costs or monitoring costs beyond those associated with in-house provision, plus the firm may have staff who can do a better job of teaching the relevant material. Second, the available educational institutions simply may not offer the relevant training, whether because of gaps in their faculty or because of insufficient demand to justify paying the fixed costs required to do so. Continuing down this line of thought, if schools find it hard to offer a particular curriculum because of the thinness of the market, workers may find the portability of their training correspondingly limited. At a minimum, workers may face higher search costs when trying to locate firms that value their "general" training. These search costs generate a degree of monopsony power that firms may exploit, an insight to which we return in Sections 5.4 and 5.6.

5.2 Meanwhile, on the empirical side

5.2.1 Tenure and wages

Beckerian theory clearly predicts that sharing of the returns to specific training should result in earnings of workers who remain at their firms being higher than workers with the same level of experience, but who have left their firm for alternative employment. While several studies had confirmed a positive cross-sectional correlation between workers' wages and their tenure at a particular firm, economists remained suspicious about whether this finding represented a causal impact of tenure, working through the accumulation of firm-specific human capital, or whether it represented an artifact of dynamic selection on match quality (or some of both). The idea of heterogeneous match quality in employment relationships dates back to the seminal work of Johnson (1978) and Jovanovic (1979). Mincer and Jovanovic (1981) and Heckman (1981) note that heterogeneity in match quality could bias estimates of the impact of tenure on wages. To distinguish the relative importance of the two explanations for the cross-sectional pattern requires a relatively long panel data set with carefully measured job tenure and earnings.

Altonji and Shakotko (AS, henceforth) (1987) use data from the 1968 to 1981 waves of the Panel Study of Income Dynamics (PSID). The PSID provides precisely the type of high-quality panel data needed to examine the impact of tenure on wages. To break the correlation between unobserved match quality and tenure, AS create an interesting instrument: the deviation of tenure from its mean.[59] Thus, if we let $T_{i,j,t}$ denote the tenure of the i-th worker at the j-th job at time t, AS define $\tilde{T}_{i,j,t} = T_{i,j,t} - \bar{T}_{i,j}$ where $\bar{T}_{i,j}$ is the mean tenure of the i-th worker at the j-th job over the sample period. As AS (1987, 440) argue, "$\tilde{T}_{i,j,t}$ is orthogonal by construction to [the fixed job match and individual components], which are constant during job j and which embody permanent individual and job heterogeneity. If $\tilde{T}_{i,j,t}$ is also uncorrelated with the transitory error component ... as we maintain in interpreting the empirical results, then it is a valid instrumental variable."

Equipped with their interesting instruments, AS estimate that OLS substantially overstates the return to tenure. Evaluating the differences at 10 years of tenure, AS find that OLS overstates the return to tenure by a factor of 10, although subsequent adjustments to their estimates reduce this differential to a factor of five. While AS (1987, 454) note that the "many caveats" that attend their study could imply a bias somewhat smaller than their preferred point estimate, they nonetheless argue that "it is clear that heterogeneity is responsible for the much larger least squares estimates of the tenure profile which have appeared in the literature. The evidence indicates that the positive bias in the least squares estimates arises from both individual heterogeneity

[59]In their specification, AS use three measures of tenure: tenure, tenure squared, and an indicator variable for being beyond the first year of the job. They employ the same deviational strategy to create instruments for all three measures. We remind more youthful readers that AS (and the other papers described here) write in the context of a common coefficient worldview. Reconsidering this particular instrumental variables strategy in a world of heterogeneous coefficients on tenure would add value.

and job match heterogeneity."

In a closely related paper, Abraham and Farber (AF, henceforth) (1987) also use data from the 1968 to 1981 waves of the PSID. Unlike AS, AF confine their study to two narrower populations: non-union managerial and professional workers and non-union blue-collar workers. While AS conditioned on union status, AF argue that the contracting process differs enough for union workers to justify a focus solely on non-union workers.

AF face the same daunting endogeneity problem as AS (and everyone else). They observe that higher unobserved match quality implies, on average, a longer completed job duration. Building on this idea, they exploit the panel nature of the data and estimate a Weibull proportional hazard model using the available data on job spells, and incorporate right censoring of spells still in progress in the 1981 wave. Using the observed completed duration for the observations that have it, and the predicted expected duration for the censored observations, they construct residuals from a regression of tenure on completed duration and use these as instruments. Intuitively, the residuals proxy for the unexpected component of job duration and thereby for unobserved match quality. They include experience, experience squared and tenure in their log wage equation, and use pre-spell experience, pre-spell experience squared, and the estimated tenure residual as instruments.[60] The IV and OLS estimates differ very little for experience and experience squared. In contrast, the tenure coefficient falls by about 80 percent when moving from OLS to IV for the blue collar workers and by about 40 percent for the managers and professionals.

Topel (1986) offers a similar analysis. He uses the Longitudinal Employee-Employer Data (LEED) from the US Social Security Administration. These data match employee

[60] They also implement a "control function" version of their identification strategy that includes the completed duration on the right-hand-side of a wage equation estimated by OLS. By construction, the predicted job durations of the censored observations are a nonlinear function of variables that appear in the earnings equation. Put differently, identification for these observations arises solely from the nonlinear functional form, as in a bivariate normal selection model with no exclusion restriction. Though less obvious, the same point applies to the strategy that uses the tenure residuals as instruments for tenure. While reliance on the non-linear functional form does not *necessarily* render these strategies invalid, it does require some optimism on the part of the researcher.

quarterly earnings reports to their employer from 1958 to 1972. While these data obviously offer superb measures of earnings for formal sector workers, they lack information on hours or weeks worked and contain only very limited worker characteristics, specifically age, race and sex. Topel selects workers born after 1938 so that the oldest worker is 18 when the panel begins. He further limits the data to white men with strong labor force attachment.

Topel constructs a job matching model that incorporates workers searching to find better matches as well as general and job-specific human capital. His empirical setup, like that in AF, uses the expected duration of employment spells to control for the selective job changes (and, therefore, selective retention of workers). Like AF, Topel finds substantial returns to experience, but virtually no return to tenure. In contrast to AF, he finds that job matching accounts for about 25 percent of the returns to labor market experience.

Collectively, these papers indicate little role for firm-specific human capital, at least in the context of the Beckerian model. While these papers find substantial returns to labor market experience, these returns appear common across all firms. Put differently, the combination of big returns to experience and small returns to tenure implies that workers mainly receive training with value outside of one specific firm.

In his (1991) paper, Topel reverses himself and argues that wages do increase with tenure. He offers two pieces of evidence. First, he documents that workers with greater tenure experience larger average wage losses when displaced in the 1984 and 1986 waves of the U.S. Displaced Workers Survey (DWS). As Topel notes, though consistent with the empirical importance of firm-specific skills that accumulate with worker tenure, jobs that last longer may pay permanently higher wages for other reasons as well. He argues that distinguishing among alternative explanations for this pattern requires longitudinal data on the job histories of individual workers.

Thus, Topel's second piece of evidence builds on data from the first 16 waves of the PSID (i.e. from 1968 to 1983, and covering labor market outcomes from 1967 to 1982). Topel restricts his analysis to white men ages 18 to 60 who were not self-employed,

employed in agriculture, or employed by government. The logarithm of average hourly earnings, deflated using a wage index constructed by Murphy and Welch (1992), serves as the outcome measure of interest. With these data, he estimates an empirical model of the form

$$y_{i,j,t} = X_{i,j,t}\beta_1 + T_{i,j,t}\beta_2 + \epsilon_{i,j,t} \tag{19}$$

where i indexes the individual, j indexes the job, t indexes calendar time, $X_{i,j,t}$ is total labor market experience, $T_{i,j,t}$ is current job tenure, and (β_1, β_2) are parameters to be estimated. In this model, β_2 represents the "structural" tenure effect; when $\beta_2 > 0$, workers will not change jobs unless their new wage compensates for their foregone returns to tenure at their old firms.

In models like (19), we worry about the correlation of experience and tenure with the error terms. To aid in thinking about such correlations, Topel decomposes the error as

$$\epsilon_{i,j,t} = \phi_{i,j,t} + \mu_i + \nu_{i,j,t} \tag{20}$$

where $\nu_{i,j,t}$ combines market-level shocks and measurement error, μ_i is an individual fixed effect, and $\phi_{i,j,t}$ is the match quality effect. Notably, the match effect depends on time in Topel's model. He specifies an auxiliary regression to model the evolution of $\phi_{i,j,t}$ as

$$\phi_{i,j,t} = X_{i,j,t}b_1 + T_{i,j,t}b_2 + u_{i,j,t} \tag{21}$$

When estimating the model given in (19) with OLS, $\mathrm{E}(\hat{\beta}_1) = \beta_1 + b_1$ and $\mathrm{E}(\hat{\beta}_2) = \beta_2 + b_2$. Thus, whether OLS underestimates or overestimates the returns to tenure depends on the sign of b_2.

Adding a time subscript to the work-job match component $\phi_{i,j,t}$ renders the AS instruments invalid. In their place, Topel develops a two-step estimation strategy that relies heavily on the longitudinal nature of the data and allows him to estimate (β_1, β_2) and $b_1 + b_2$. Applying his strategy to the PSID data, he finds strong evidence that

$b_1 + b_2 \neq 0$, implying empirically important time-varying match effects in equation (21). He also estimates cumulative returns to tenure over the worker's first 10 years in a job of about 33 percent.

As Topel notes, these results concord with the losses of displaced workers estimated using the Displaced Worker Survey (DWS), to which we turn shortly. At the same time, they well exceed the tenure effects found in AS (1987) and AF (1987), and, for that matter, Topel (1986). Both Topel and subsequent responses by Altonji and Williams (1997, 2005) discuss many reasons for these differences. While a full discussion of these reasons would take us too far afield, we highlight three that give the flavor of the discussion and also illustrate just how much care the early participants in this literature took with their data:

1. Measurement error in the PSID measure of job tenure was pretty severe. First, some of the tenure data are reported in intervals. More generally, Topel (1991, 150) remarks on the "measurement error that is known to plague survey data" like that from the PSID. For example, Topel notes that 3.8 percent of workers in his PSID sample who do not change jobs report reductions in tenure while many others report implausibly large increases.[61]

2. The use of time trends versus the Murphy and Welch (1992) wage index to deflate wages has a sizable effect on the estimates that varies with the method used to handle attrition from the PSID panel.[62] Many researchers may find this a shocking finding – how can such an innocuous decision as how to adjust wages for economy-wide trends matter so much?[63] As Altonji and Williams (2005, 386) note, "Which detrending procedure is more robust with respect to the potential for bias from attrition on the basis of time-varying error components is an open

[61]Early waves of the PSID also had problems related to connecting job characteristics across years; see e.g. Kambourov and Manovskii (2009).

[62]When applied researchers use panel data, they often ignore its potential to induce bias. This hopeful view has at best mixed support in the literature; for the PSID see e.g. Becketti et al. (1998) and Fitzgerald et al. (1998).

[63]See Sacerdote (2017) for a stunning recent example of the substantive importance of choices around inflation adjustment.

question. It might be best to try to model such attrition ...”

3. Because of the design of the PSID survey instrument, researchers had to rely on “average hourly earnings last year” as their wage measure. Moreover, the PSID measures job characteristics as of the current survey but asks about earnings for the previous calendar year. In response, AS take the wage measure from the survey in year t and other characteristics from the survey in $t-1$ while Topel takes all information from the survey in year t, but excludes observations with less than a year of tenure. As Altonji and Williams note, neither procedure corresponds exactly to the model; as such, both could result in biased estimates.

Reading these papers today leaves (at least) two strong impressions. First, the authors took extreme care with their data. They worried about the quality of the measures their data provided, they worried about the limitations inherent in their data, and they worried about the correspondence between the data and the conceptual objects in their models. Second, and related to the last point, theory played an important role in these papers in guiding the specification of empirical models and in interpreting the resulting estimates.

Looking back at these papers, most written over 30 years ago, some readers may (not unreasonably) want to quibble about their choice of instruments or to complain about other applied econometric sins. In our view, such criticisms have their place in a weighting of the evidence and in designing a replication or update, but they miss the most important point. These papers represent impressive attempts to test the available models of firm-provided training and to learn about the importance of firm-specific and general training using data *without* any direct measures of training. As discussed in Section 4, with the advantage of hindsight, we now know that jobs differ substantially in the amount of on-the-job training provided. This makes inference on the degree of specificity through the use of common experience-earnings and tenure-earnings profiles a virtual impossibility.

Another potential criticism of these early papers builds on concerns about the endogeneity of job duration and thus tenure in empirical models like that in (19).

Workers and firms may choose if and when they dissolve their job match based in part on factors unobserved by the researcher but correlated with both wages and tenure, and may do so in ways resistant to some or all of the identification strategies adopted in these papers. One thread of the related literature sees worker job shopping ("voluntary transitions") as problematic, while seeing layoffs and firings by the firm ("involuntary transitions") as relatively benign.[64]

In the spirit of that literature, what do the data have to say about the relative prevalence of the two types of separations? Barron and Loewenstein (1985) provide an analysis of why firms believe turnover occurred using the 1978 Employment Opportunities Pilot Project (EOPP) data. In this survey, firms reported that 2.93 workers per 100 quit, 1.62 workers per 100 were laid off, and 0.91 workers per 100 were fired for cause, for an aggregate turnover rate of 5.46 per 100 workers each month. Taken at face value, then, 46 percent of workers who leave did so at the firm's request, implying a roughly equal mix of voluntary and involuntary transitions.

These statistics come with some caveats. First, the EOPP data suggest that firms sometimes induce workers to quit (0.29 workers per 100 per month) and sometimes lay workers off for cause (0.36 workers per 100 per month). Presumably workers sometimes induce firms to let them go as well. Second, the EOPP represents a random sample of establishments, not of workers, and so over-represents workers in small establishments relative to their share of employment. Third, we would expect the share of voluntary and involuntary separations to vary with local economic conditions, with the aggregate business cycle, and over the lifecycle of individual workers. This variation makes inference about the importance of firm-specific human capital even more difficult. None of these caveats undermines the first-order qualitative point that a substantial fraction of the workers analyzed in the papers above experience potentially problematic voluntary transitions. That point led researchers to study narrower samples of displaced workers.

The rapid decline in American manufacturing employment during the Reagan years

[64]Some other ways of looking at the problem, including various models in the search and matching literature, imply no substantive distinction between quits and layoffs.

provides a set of "natural experiments" wherein large groups of workers lost their jobs for plausibly exogenous reasons. The U.S. Displaced Worker Survey (DWS) includes people who report having lost a job in the last five years due to a plant closure, an employer going out of business, or a layoff without recall. Kletzer (1989), Topel (1990, 1991), Carrington (1993), and Neal (1995), among others, use the DWS to examine how the wage loss associated with displacement varies with worker tenure. These papers all find (1) very large average wage reductions associated with displacement and (2) larger average reductions for workers with longer tenure at their former firm. Firm-specific training may play an important role in explaining both patterns.

There are, however, reasons for caution in regard to the DWS data. First, Topel (1990) and Jacobsen et al. (1993) provide evidence that the retrospective nature of the DWS data leads to under-reporting of job losses that increases with the temporal distance from the interview date. Second, as Barron and Loewenstein (1985) demonstrate, mass layoffs may hide some (less plausibly exogenous) layoffs for cause if firms combine a set of troublesome workers into their own mass layoff, or find ways to over-represent such workers in a broader layoff. Third, the DWS relies on self-reports for both current earnings and earnings on the job pre-displacement job. For those of us who have difficulty remembering what we had for dinner last night, this seems heroic.[65] Moreover, while the DWS provides a sample of workers who experience plausibly exogenous job loss, there is for that very reason no pretense that it represents a random sample of all workers in the U.S. economy.

In a very influential paper, Jacobson et al. (1993) (JLS, henceforth) use administrative data from Pennsylvania to examine the impact of mass layoffs and plant closures on worker earnings. The data contain quarterly earnings records for five percent of Pennsylvania workers from 1974 through 1986. Within their data, JLS identify workers displaced in a mass layoff or plant closure. Their data also encompass a large sample of non-displaced workers, giving their analysis the look and feel of a modern

[65] Apparently, the BLS and the Census Bureau had concerns about recall bias as well, as they reduced the DWS recall period from five to three years starting with the February 1994 survey; see Gardner (1995).

program evaluation paper.[66]

JLS document extremely large average earnings losses associated with displacement. Indeed, they show that average earnings begin to fall three quarters before displacement, with losses exceeding $1,000 in the quarter immediately prior to displacement. Average losses exceed $3,000 in the first quarter after displacement and slowly adjust to about half of that three years of after displacement, at which point they stabilize.[67] While admitting their data do not allow them to distinguish among firm-specific human capital, internal labor market issues, or other sources of rents, they do examine the earnings of workers that remain in the same industry. JLS (1993, 706) conclude that "our results indicate that there is something intrinsic to the employment relationship itself that is lost when workers are displaced. If it is workers' skills that are lost, these skills must be firm-specific, as opposed to industry-specific ... [O]ur finding that losses are large for almost every group we studied suggests that wage premiums—whether due to firm-specific skills or to internal labor markets—must be commonplace."

JLS were rightfully cautious about ascribing the substantial earnings losses upon displacement to specific human capital. Like the data in all of the papers considered up to this point, their data contain no direct measures of employer training. In the next subsection, we consider what the literature has learned from data that do contain such measures.

5.2.2 Does training lower starting wages?

The PSID became the first major data set to try to measure firm-provided training. As noted in Table 1, in 1976 and 1978, the PSID asked: "On a job like yours, how long would it take the average new person to become fully trained and qualified?"

[66] Like the LEEDS data used by Topel (1986), the Pennsylvania data lacks information on hours worked, weeks worked, education, and occupation. It does include information on industry.

[67] JLS estimate the average effect of displacement today versus no displacement during the period covered by the data. Krolikowski (2018) shows that estimating the effect of displacement today versus no displacement today but maybe displacement in the future meaningfully reduces the average effect. His parameter has the virtue of not conditioning on future outcomes and not varying with the length of the panel. Recent work, e.g. Jahromi and Callaway (2022), has examined heterogeneity in the earnings response to displacement. That line of work shows that the mean effect conceals quite substantial heterogeneity, with some workers actually better off due to displacement.

This formulation has several notable features. On the positive side, asking only a single question, rather than a battery of questions that might take several minutes, should reduce item non-response and increase overall survey completion. Second, it makes no attempt to describe the training the worker receives. In addition to not collecting any information about duration, or content, or provider, this formulation fails to differentiate between general skills and specific skills. Indeed, a respondent could easily interpret this question as asking how much experience—as distinct from actual training—it takes to become fully qualified. Finally, this single, relatively short question implicitly imposes a non-trivial cognitive burden on the respondent, who must produce an internal definition of "fully trained and qualified" and then take some implicit weighted average of their own experience and that of co-workers in the same job. Can workers provide informative responses to it?

Brown (1989b) made the first attempt to relate the PSID training measure to wages. He limits his analysis to PSID household heads (mostly men) who responded to the 1976 training question and who remained in their 1976 position for at least two interviews.[68] The mean response to the PSID training question in Brown's sample equals 1.10 years, with a standard deviation of 1.56 years. Thus, this measure of training exhibits both substantial heterogeneity and a skewed distribution with a long right tail, given the impossibility of responses below zero.

Brown takes the PSID measure at face value and goes looking in the data for different empirical returns to tenure before and after the time when the worker becomes "fully trained and qualified" (FTQ). He expects, based on Beckerian theory, that wages should increase rapidly during this period of intense training period and less rapidly thereafter. He finds exactly the expected pattern. Evaluated at the mean time to FTQ of 1.1 years, the wages of FTQ workers exceed their wage at hiring by about 20%. Put differently, workers in jobs where it takes 1.1 years to become FTQ see their wages grow about 20% relative to workers in jobs simple enough that workers start out FTQ.

[68] The sample includes the PSID's SEO subsample and so over-represents low-income households. The published paper makes no mention of sampling weights.

That wages increase very quickly as workers become FTQ indicates that the PSID training measure effectively captures aspects of worker skill accumulation. More generally, it confirms a strong positive correlation between wages and training. At the same time, the PSID data impose important limitations on the ability of Brown's analysis to fully engage with Beckerian theory. First, as noted, the PSID question does not attempt to measure the firm-specificity of the training. This renders the 20% figure difficult to interpret as more than just evidence that training matters for wages, though even that qualitative finding made a splash at the time. Second, as noted, the PSID does not collect information on training costs, either directly or in the form of proxies such as duration. Thus, it cannot support tests of the aspects of the Becker model regarding who pays for training. Third, the absence of questions concerning ongoing training, whether to maintain existing skills or build new ones when technology changes or promotions occur, rules out an examination of any relationship between training and wages at post-FTQ tenures.

Back in the late 1970s and early 1980s, the US Department of Labor constructed a novel data set: the Employment Opportunities Pilot Project (still EOPP). We follow Barron et al. (1989) (BBL, henceforth) in focusing on the second (1982) EOPP wave as it provides the most detailed information about training. As noted in Sections 3.3.1 and 4.2.2, this data set sampled both workers and firms from the same geographic areas and so allowed researchers to study the search behavior of workers and firms in the same local labor markets.

The EOPP employer survey asked about the last worker hired at least three months prior to the date of the interview and the training offered to these workers. BBL (1989) report on five distinct (and mutually exclusive) types of training: formal training offered by specially trained personnel, formal training offered as a part of job orientation, informal training offered by supervisors, informal training offered by coworkers, and time spent by the newly hired worker watching others work rather than working themselves. The sample mean of the sum of these five measures equals 151.1 hours, with a standard deviation of 206.8 hours. This represents a pretty substantial investment. For workers employed 40 hours per week, 151.1 hours equals about 29 percent of the totals hours worked in the first three months of employment. That the standard deviation substantially exceeds the mean implies a long right tail in the distribution of training hours.

The individual training type measures provide valuable information as well. Mean formal training equals 10.9 hours, with a standard deviation of 51.9 hours. Orientation consumes an average of about 5.9 hours, with a standard deviation of 13.3 hours. Informal training by supervisors averages 54.3 hours with a standard deviation of 93.2 hours, while informal training by coworkers averages 26.9 hours with a standard deviation of 63.5. Finally, the amount of time the worker spends watching others averages 53.1 hours, with a standard deviation of 100.4. Three training types have standard deviations at least twice their means, implying long right tails for their hours distributions. The impressive amount of heterogeneity overall suggests that at least some firms specialize in particular training types, supplying many hours of one or two training types and few if any hours of the others.[69]

In addition to training, the EOPP also asked employers about search activities related to the hiring of the most recent worker, the worker's starting wage, the worker's productivity growth in their first three months at the firm, and the worker's wage at the end of that period.[70] Thus, these data provide a unique opportunity to test the Beckerian model because they contain both the starting wage and a second wage measured after the worker completes some (if not most or even all) of their initial training.

BBL (1989) provide this test. They find no evidence that workers bear the cost of training: the correlation between starting wages and hours of training in the first three months of employment proves neither statistically significant nor substantively important. In contrast, training hours do predict both productivity growth and wage

[69]The data allow investigation of the joint distribution of hours devoted to the specific training types but none of the related papers describe the joint distribution in detail.

[70]The productivity growth estimate comes from a question that asks employers to rate the productivity of the reference worker during their first two weeks on the job on a scale of 1 to 100. See BBB (1997b, 132-135) for more detail.

growth, although the coefficient on training in the productivity growth equation is five times as large as that in the wage growth equation.[71]

BBL (1989) further document that firms see more applicants and spend more time per applicant when hiring workers for positions that receive more training hours. We interpret this substantively and statistically significant relationship as indicating that firms want to focus their training investments on relatively more talented workers. Why would firms spend more time searching for and evaluating applicants they expect to receive more training? We think firms believe (correctly) that worker ability complements firm training and that firms understand that they help pay for and thus share the benefits of investments in on-the-job training.

Finally, BBL (1989) observe that this matching process presents a problem for ascertaining the impact of training on starting wages. Merely comparing the starting wages associated with positions with more or less training will fail due to the deliberate sorting of workers with greater abilities into jobs that offer more training. While researchers may observe some of the relevant dimensions of worker ability on which firms sort (e.g. schooling, experience, etc.), they most likely do not observe all of them. In our view, this fundamental issue makes finding truly exogenous empirical variation in training incidence problematic indeed.

BBB (1999) use the same 1982 EOPP survey data as BBL (1989) along with the (deliberately very) similar data from the 1992 Small Business Administration (SBA) survey, which we described in Section 4.2.2. They present three analyses using (where possible) both data sets. The first two analyses test the human capital model prediction that workers who receive general training should pay for it (at least in part) via lower starting wages. In a standard log starting wage model, they find some modest evidence for this proposition but only when including one or two proxies for worker ability. The first proxy they call "job complexity"; it consists of the employer response to a question regarding how long it takes to become proficient on the focal worker's job. The second proxy consists of the employer response to a question about how much

[71] Training has a tiny, negative, and statistically insignificant coefficient in the starting wage equation.

effort they devoted to screening workers for the job, with the idea that more effort will yield (otherwise unobservably) more able workers. In the second analysis, they use responses to questions about whether the focal worker received more, less, or the same training as other workers in the same job, and more, less, or the same pay as other workers in the same job and obtain a somewhat puzzling asymmetric finding in which workers who receive less training tend to receive higher pay, but those who receive more training do not tend to receive lower pay.

The third analysis builds on a simple human capital model that predicts that initial productivity growth should exceed initial wage growth for workers receiving training. The authors estimate parallel linear regressions of log wage growth and log productivity growth (both measured over the first two years on the job) on total on-site training hours in the first three months on the job (and total off-site hours in the SBA data) and various conditioning variables. The conditioning set does not include training received after the first quarter on the job, and the analysis structure limits the sample to workers hired at least two years prior to the interview for jobs into which the firm hired no one else in the interim. With those caveats, the estimates strongly support the simple human capital model for on-site training, and thus provide indirect evidence of workers covering part of the cost of firm-provided general training.

5.3 Training in imperfect labor markets

At least since the publication of Stevens (1994), economists have attempted to explain why firms might pay for general training, and why firms' paying for training could result in sub-optimal training levels. Before reviewing that paper and its intellectual progeny, we want to adjust the reader's expectations downward by noting the limited amount of guidance this literature provides regarding whether the distortions arising from imperfect labor markets—what Acemoglu and Pischke (1999b) refer to as labor markets with frictions—result in substantial inefficiencies or not. Ideally, our models, when taken to the data, would allow us to calculate the optimal level of training, to quantify the departure from that optimal level, and to decompose the difference into components due to specific frictions. Instead, the literature offers something much more modest, but still quite useful, namely a set of comparative static results regarding the signs of the effects of specific frictions on the difference between realized training levels and optimal training levels.

The current state of play prompts a couple of reactions. First, rich opportunities remain for innovative applied theory work in this domain. Second, given the measurement issues lovingly described in Section 3, one can understand the general reluctance to "discipline" (as those kinky macroeconomists say) the models with the data, though we think such work would have substantial value at the current margin. Such work could test comparative static predictions, as we do in Section 5.4, or could seek to quantify the role of specific frictions in specific contexts. Both types of studies would increase our understanding of how firm training works and would improve qualitative and quantitative policy analysis. Overall, the existing literature will disappoint readers expecting iron-clad (or even chocolate-coated) justifications for particular policy prescriptions. We hope that our presentation will serve to transform such disappointment into research action.

We begin our tour of the literature with a discussion of Stevens (1994), who poses a question at once very basic and surprisingly deep: How might we alter the Beckerian taxonomy to allow for a more nuanced view of how the market values firm training? To understand how Stevens answers her own question, imagine a labor market with $n+1$ firms.[72] The worker's incumbent firm, which we label firm zero, provides training in the first period. In the subsequent period, the vector v describes the worker's productivity at every firm, where

$$v = \{v_0, v_1, v_2, \cdots, v_n\} \tag{22}$$

In Stevens' framework, the Beckerian notion of general training transforms (22) into

$$v = \{g, g, g, \cdots, g\}, \tag{23}$$

[72]Stevens treats the number of firms as a parameter, but it is straightforward to make it endogenous.

with g constant across firms. Incorporating specific training, (22) becomes

$$v = \{g + s, g, g, \cdots, g\}, \tag{24}$$

where s reflects the value of the firm-specific component of the training. The configurations (23) and (24) capture just two possible ways in which the market may value training. For her model, Stevens assumes

$$v = \{g + s + \epsilon_0, g + \epsilon_1, g + \epsilon_2, \cdots, g + \epsilon_n\} \tag{25}$$

where ϵ_j $j \in \{0, \cdots, n\}$ represent random shocks to the value of training, revealed as private information to each firm once workers receive their training. Thus, as in equation (22), each firm places a unique value on the training of each worker. The specific training s gives incumbent firms an advantage in retaining their workers, but the random shocks provide opportunities for competing firms to bid workers away from their incumbent firms.

Stevens assumes Bertrand competition for workers' services. Thus, upon realization of the shocks, the workers essentially hold an auction for their labor services. Put differently, the absence of search frictions in the model means that workers simply choose to work at the firm with the highest valuation of their services, albeit at a wage equal to the second highest valuation.

Several features of the (partial) equilibrium warrant mention. Returns to non-incumbent firms decline as the amount of investment in specific training increases because the greater value of the specific training helps the incumbent firms retain their trained workers. This mechanism provides an incentive for firms to increase the provision of specific training, perhaps beyond the socially optimal level. In contrast, workers fully reap the benefit to general training once we have at least one competing firm. Because general training is valued at each firm in the market, Bertrand competition ensures that increases in general training result in an increase in the wages that workers receive.

Stevens emphasizes the role of imperfect competition, so a natural comparative static asks what happens when we increase the number of firms, n. In her model, an increase in the number of firms reduces profits for both incumbent firms and non-incumbent firms. It also increases the likelihood that workers leave their incumbent firm because each additional ϵ_j draw provides another chance for a non-incumbent firm to value a worker's services more than their incumbent firm. Increases in n raise workers' returns to specific training but reduce firms' return to specific training due to the increase in turnover.[73]

This paper provides valuable insights into how the structure of the labor market affects the allocations of general and specific training. As always, the theoretical results necessarily follow from the model assumptions. For instance, the assumption of Bertrand competition ensures that workers receive the full benefits of general training. The next set of papers we consider strives to understand how features of the labor market affect workers' returns to general training and induce firms to pay at least a portion of the general training costs by exploring alternative models.

Acemoglu (1997), Acemoglu and Pischke (1998), and Acemoglu and Pischke (1999a) (AP, henceforth) collectively provide a detailed analysis of how various features of the labor market might affect the amount of training and the allocation of the costs and benefits of training. Our exposition will follow AP (1999a) as it provides the most extensive theoretical discussion of the issues. AP begin by assuming a two-period model. They normalize worker productivity to zero in the first, early career, period, in order to simplify the math. During period one, any general training (τ) occurs at a cost $c(\tau)$. The cost of training is assumed to satisfy Inada-like conditions $c(0) = c'(0) = 0$, $c''(\tau) > 0$, and $\lim_{\tau \to \infty} c'(\tau) = \infty$. After training, workers' produce $f(\tau)$, where $f(0) = 0$, $f'(\tau) > 0$, and $f''(\tau) < 0$. Setting the discount rate to zero, it is easy to show that the socially optimal level of training τ^* satisfies

$$c'(\tau^*) = f'(\tau^*), \tag{26}$$

[73] A sufficient condition for this result is that the distribution of ϵ's is log concave, which means that the logarithm of the density is a concave function.

i.e., it exactly balances the marginal costs and marginal benefits of training.

AP ask "How might markets provide the incentives for firms and workers to invest in the socially optimal level of training?" To begin to address this question, we need to flesh out the rest of the model. In period two, workers produce $f(\tau)$ at either the incumbent firm or a competing firm because all training is general. AP assume that employment relationships end exogenously with probability q, $0 < q < 1$. But because all training is general, workers who lose their jobs simply find employment at an alternative firm offering the same wage.

To begin their analysis, AP consider the case where workers cannot finance any training, which implies a first-period wage equal to productivity (again normalized to zero). AP note that some of the literature incorrectly argues that firms might undertake to provide general training in a context with credit-constrained workers. In fact, this cannot be an equilibrium strategy because workers' ability to command a wage $f(\tau)$ in the marketplace in the second period will thwart any attempt by firms to recoup such investments. With no way to realize a return on investments in general training, profit-maximizing firms have no choice but to offer no training.

In contrast, the fully competitive model without constraints on the financing of training yields fundamentally different allocations of training and of the costs and benefits of that training. In this version of the model, wages in the second period again equal $f(\tau)$. Given that firms must pay a wage equal to the full marginal product in the second period, they will prove unwilling to finance any general training in the first period. Workers, on the other hand, have every incentive to invest in the general training. As they pay the full cost and reap the full benefits, equation (26) guides their investment decisions and they achieve the first-best outcome, just as in Becker's world. But AP want to discover conditions that will induce firms to pay for general training. To model how this might happen, AP distinguish between the second-period wage offered at the incumbent firm, $w(\tau)$, and that offered at alternative firms, $v(\tau)$, and make the critical assumption that $v(\tau) < f(\tau)$. AP (1999a, 542) then consider why $v(\tau) < f(\tau)$ might hold:

> We show that a range of plausible frictions, such as search, informational asymmetries, and efficiency wages, lead to this type of distortion. Furthermore, even when the labor market is frictionless, complementarities between technologically general and specific skills may induce firms to invest in the general skills of their workers. Finally, we also show that labor market institutions such as union wage setting and minimum wages, which also compress the structure of wages, may encourage firms to invest in the general skills of their employees. Therefore, our model predicts that in a variety of circumstances, we should observe firm investments in general training.

In Section 3 of their paper, AP discuss six specific explanations for why $f(\tau)$ might exceed $v(\tau)$: search and monopsony power, asymmetric information, firm-specific human capital, efficiency wages, wage floors such as minimum wages, and unions. Three of the six represent "old news" in the literature, namely firm-specific training, minimum wages, and unions, the last two going back to (at least) Mincer and Leighton (1980). Thus, we engage only with the remaining three explanations here.

Stevens (1994) worried about monopsony power resulting from specific training and from firms' heterogeneous valuations of general training. In contrast, AP worry about monopsony power arising via search costs, which they term "search-induced monopsony." The literature has long recognized that sequential search provides firms with some monopsony power; see, for instance, Albrecht and Jovanovic (1986). AP contribute by pointing out that this monopsony power provides an incentive for firms to invest in general human capital. They have in mind two ways that search lowers the worker's return from changing firms.

First, job changes may imply a (costly) period of unemployment. Letting c_u denote the costs of changing jobs, workers will move only if $f(\tau) < c_u + v(\tau)$, and so may not move even if $v(\tau)$ slightly exceeds $f(\tau)$. Yet many employee-initiated job changes result in direct job-to-job transitions with no intervening spell of unemployment.[74]

[74] Manning (2003) and Nagypál (2005) note that half of new hires arrive directly from other employers; see Manning (2011) for a discussion. As these authors know, at the margin, measurement issues abound

If the firm is exploiting a quasi-rent and paying less than the value of the worker's marginal product, we expect the employee to initiate the turnover.

Second, the presence of search costs can provide a degree of monopsony power as outlined in Albrecht and Jovanovic (1986) and many other search models. The monopsony power in these models arises from the sequential nature of search. But at least since the publication of Morgan and Manning (1985), economists have recognized that the nature of the labor market determines whether workers search sequentially or simultaneously. Under simultaneous search, workers with multiple offers may hold an "auction" for their services, much as in Stevens (1994). Does the potential auction solve the monopsony problem? Perhaps, but we are dubious.

The problem has a firm side as well. Firms face a hold-up problem because they spend real resources recruiting, hiring, and training workers. When workers quit, firms must replace them and, in doing so, incur such costs yet again. Table 2 of Manning (2011) provides estimates of these quite substantial costs. So which side has the "hold-up" power in employment relationships? Perhaps the answer varies over time depending on the state of the labor market. As we write in 2023, many employers in the U.S. offer signing bonuses and complain about "shortages" of workers. Workers with offers from competing firms have a lot of leverage in this economic environment. In contrast, back in the beginning of 2010 when the unemployment rate sat at 10 percent (rather than below four percent), workers had trouble obtaining outside offers and firms replaced departing workers with ease. This line of reasoning suggests variation in the extent of firm-provided general training over the business cycle. We will return to these issues in Section 5.6.

AP also consider the implications of efficiency wage models for investments in general training. These models seek to provide a theory of involuntary unemployment

when attempting to distinguish job-to-job transitions from job-to-unemployment-to-job transitions. For instance, if a worker has a spell of non-employment lasting one or two weeks between jobs, did they make a job-to-job transition with a vacation built in or did they get laid off but found a job very quickly once unemployed? In the past few years, the rapid increase in internet job search has facilitated a large expansion of on-the-job search, further diminishing foregone earnings as a component of search costs. Of course, a job-to-job transition may still require substantial search costs as workers must juggle their current jobs with the necessity of interviewing for alternative jobs and evaluating any alternative offers.

predicated on the assumption that firms have trouble detecting worker shirking. As a result, firms pay wage premia to workers, which create a quasi-rent. Workers then avoid shirking (or at least shirk less) in order to avoid detection, firing, and loss of the quasi-rent. AP show that in this type of model firms will usually pay for some general training. At the same time, they note that these models do not generate the poaching of trained workers from other firms. Given the empirical importance job-to-job transitions, this represents a serious drawback to the application of efficiency wage models to the division of the costs and benefits of training and their relationship to job turnover.

Asymmetric information provides an interesting possibility to explain firms' willingness to invest in general training, and AP focus on this explanation in their 1998 *Quarterly Journal of Economics* paper. The central idea behind models of asymmetric information in the labor market is that workers differ in their abilities and that employers that hire them learn more about their ability than employers that do not. As a result, when a worker looks for a new job, an information asymmetry exists between the incumbent employer and alternative employers. AP add training to this standard story. To see the intuition, consider, from the point of view of alternative firms, whether to make second-period offers to workers trained by their incumbent firms in the first period. As incumbent firms have better information, they may match or beat offers of alternative firms for workers they know to be high ability and say good riddance to workers they know to have low ability. The alternative firms then end up hiring only the low ability workers, a phenomenon the auction theory literature calls the "winner's curse." In extreme cases—such as Akerlof's (1970) famous "market for lemons" paper—the adverse selection associated with the winner's curse can eliminate a market completely, though AP set up their model to avoid this equilibrium.[75]

Let us formalize AP's ideas a bit. Let η_i be the productivity of the i^{th} worker. With probability p, the worker has low productivity, normalized to zero ($\eta_i = 0$) and

[75]Katz and Ziderman (1990) and Chang and Wang (1996) also note that if the market cannot ascertain the quality of the training that a firm provides, it may have a lower market value at other firms. In a sense, objectively general training may not become subjectively general training until the market fully recognizes its quality.

with probability $(1-p)$ the worker has high productivity, normalized to one ($\eta_i = 1$). Incumbent firms must make training decisions prior to learning the productivity of the worker. For the moment, let $\hat{\tau}$ denote the level of training that firms choose; we will return to its determination in due time.

During period one, incumbent firms train their workers and learn their productivities. In the second period, workers' productivity equals $f(\hat{\tau}, \eta_i) = \eta_i * \hat{\tau}$.[76] Therefore, as $f(\hat{\tau}, 0) = 0$, incumbent firms offer low-productivity workers a wage of zero, or $w(\hat{\tau}|\eta_i = 0) = 0$. High productivity workers produce $f(\hat{\tau}, 1) = \hat{\tau}$. Call their wage offer from alternative firms $v(\hat{\tau})$ and observe that incumbent firms will optimally match the outside offer for high-productivity workers, so that $w(\hat{\tau}|\eta_i = 1) = v(\hat{\tau})$. Assume for the moment that $v(\hat{\tau}) > 0$. In this case, no high-productivity workers will leave their incumbent firms but all low productivity workers will in order to receive the higher wage on offer elsewhere. But the alternative firms, understanding that the second-period job market contains only low-productivity workers, will not offer a positive wage. Put differently, $v(\hat{\tau}) > 0$ does not yield a stable equilibrium. Instead, all workers earn zero in the second period because the high-productivity workers cannot leverage their training in the market due to the asymmetric information.

AP move the model away from this equilibrium by assuming a probability λ of exogenous separation for the high skill workers, with λ a parameter of the model.[77] Competition ensures that alternative firms will pay second period workers the expected value of their marginal products. Given the presence of some high-productivity workers in the market, this wage $v(\hat{\tau}) > 0$ and all low-productivity workers leave their incumbent firms to receive the higher wage. Thus, alternative firms pay

$$v(\tau) = \frac{\lambda(1-p)\tau}{p + \lambda(1-p)} \tag{27}$$

[76]As AP note, their model requires complementarity between ability and training. They show that if you assume a production function of $f(\hat{\tau}, \eta_i) = \eta_i + \hat{\tau}$, firms do not offer any training.

[77]AP (1998) makes this fraction endogenous by assuming that workers draw a random disutility of remaining with the firm at the end of the first period, as in Albrecht and Jovanovic (1986). The AP and Albrecht and Jovanovic models yield similar equilibria.

Incumbent firms match this wage when $\eta_i = 1$, but offer a wage of zero to the low-productivity ($\eta_i = 0$) workers, who ignore their retention offers and sign on at the alternative firm.

To determine the level of training, AP specify the following profit function

$$\pi(\tau) = (1-\lambda)(1-p)(\tau - v(\tau)) - c(\tau) \tag{28}$$

If we select the level of training to maximize profits the necessary condition is:

$$\pi'(\tau) = (1-\lambda)(1-p)\left(1 - \frac{\lambda(1-p)}{p+\lambda(1-p)}\right) - c'(\tau) = 0 \tag{29}$$

If $c(\cdot)$ is strictly convex, then the second-order condition is satisfied and we may solve equation (28) for the optimal level of training, or

$$\hat{\tau} = g\left(\frac{(1-\lambda)(1-p)p}{p+\lambda(1-p)}\right) \tag{30}$$

where $g(\cdot)$ is the inverse function of the marginal cost of training.[78] A couple of comparative statics follow immediately: If we increase p, the fraction of low-productivity workers, then firms reduce training. Similarly, if λ increases, leading to more high-productivity workers in the second-period labor market, then firms' optimal training level falls.[79]

The careful reader may notice the absence of first period productivity and wages in the profit function. As described above, AP normalize pre-training productivity to zero

[78] We know such an inverse function exists because $c(\cdot)$ is strictly convex.

[79] The same intuition leads to a variant of the AP model that predicts less training for women of childbearing age, as found in some of the papers discussed in Section 4. Such a model would parameterize the probability of job exit at the end of the first period for high ability workers separately for young women, young men, older women, and older men, e.g. $\lambda = \lambda_k$ for $k \in \{yw, ym, ow, om\}$ and then posit that λ_{yw} exceeds the other three. Barron et al. (1993) offer a similar model. At least since Polachek (1981), economists have recognized that the traditional Beckerian model embodies an incentive for occupational segregation by sex because of the costs of general training. See the fruitful debate between England (1982, 1985) and Polachek (1985a, 1985b) on this score. Kuziemko et al. (2018) provide more recent empirics, while the excellent review by Cortés and Pan (forthcoming) ponders survey evidence that women often fail to anticipate their withdrawal from the labor market. This failure may reflect underestimation of the costs of motherhood, ignorance about the preferences of a future partner, or possibly even unanticipated post-birth changes in the mother's own preferences.

and assume that first period wages equal zero. This last assumption is not innocuous. To see why, assume no discounting and suppose that firms earn positive profits, given by (28), at the optimal level of training, or $\pi(\hat{\tau}) > 0$ in notation. Suppose further that your firm offers zero wages in period one. Another firm filled with greedy capitalists (e.g. the authors of this chapter) instead offers a first period wage of $\frac{\pi(\hat{\tau})}{2}$. We will hire all the workers and make larger total profits, despite making less on each individual worker. Of course, we get our comeuppance as this bidding process will continue until $w_1 = \pi(\hat{\tau})$ and the workers receive all of the surplus.

This alteration to first period wages in the AP model does nothing to change the equilibrium level of training nor equilibrium wage levels in the second period. It does, however, affect the wage profiles of trained workers. As all workers receive a wage $v(\hat{\tau})$ in period two, wages will increase if

$$w_1 = \pi(\hat{\tau}) = \frac{(1-\lambda)(1-p)\hat{\tau}p}{p+\lambda(1-p)} - c(\hat{\tau}) < \frac{\lambda(1-p)\hat{\tau}}{p+\lambda(1-p)} = v(\hat{\tau}) \tag{31}$$

Below, we will parameterize the cost of training $c(\tau)$ and show that under reasonable parameter values wages will increase between period one and two.

Before parameterizing the cost of training, we take a second look at a couple of issues addressed in AP, but from another angle and with different lighting. To start, consider the socially optimal level of training given the constraint that a fraction p of workers do not benefit from the training. The social planner's problem maximizes output subject to the constraint that we cannot identify *ex ante* the workers who will not gain from the training. In notation, the social planner faces this problem:

$$\max_{\tau} s(\tau) = (1-p)\tau - c(\tau). \tag{32}$$

Selecting the level of training to maximize social welfare implies setting

$$s'(\tau^*) = (1-p) - c'(\tau^*) = 0. \tag{33}$$

Solving for the socially optimal level of training then yields

$$\tau^* = {c'}^{-1}(1-p) > {c'}^{-1}\left((1-\lambda)(1-p)\Big(1-\frac{\lambda(1-p)}{p+\lambda(1-p)}\Big)\right) = \hat{\tau}, \tag{34}$$

so that the competitive solution implies too little training in equilibrium. The shortfall does not result simply from having workers who do not benefit from the training. Rather, it results from non-incumbent firms not knowing which workers do and do not benefit from training. As in Akerlof's (1970) "market for lemons," the inefficiency arises because of asymmetric information in the marketplace. One can easily show that if workers can communicate their competence to alternative firms, the resulting equilibrium features the efficient level of training, with workers paid $f(\tau) = \eta_i \times \tau$ in the second period (and zero in the first period).[80]

5.3.1 Training costs and benefits under asymmetric information

Who bears the costs and reaps the benefits of training? We do not care about the accounting costs of training—clearly borne by firms—but rather about who pays the real costs and reaps the real benefits of the training. To answer this question, we compare the equilibrium outcome in the AP model with that from a model without the option to train.

The normalizations in the AP model make the case of no training easy to analyze. All workers in the economy receive the same wage (normalized to zero) in both periods, while firms earn zero profits. In contrast, the equilibrium with training results in higher wages in the initial period $w_1 > 0$ and in the second period $v(\hat{\tau}) > 0$. Thus, all workers benefit from the training, and do so equally despite the fact that some workers do not experience an increase in productivity from the training.[81]

We can also compare the no training equilibrium to the equilibrium that would

[80] While AP do not consider this case, it lies only a short theoretical distance from their discussion of training in a perfectly competitive labor market. Given the returns documented from numerous experimental evaluations of training programs, we are reasonably certain that they are not *maximizing* the social return to training.

[81] If someone outside the model economy (exogenously) paid the training costs, it would change the equilibrium only by increasing the first period wage to $w_1 + c(\hat{\tau})$.

result if the market could ascertain each worker's productivity type *after* training to learn who bears the costs of the asymmetric information. In this full information equilibrium, first-period wages would equal productivity net of training costs, or $-c(\tau^*)$ for all workers. In the second period, low-productivity workers receive a wage equal to their productivity, still normalized to zero. High-productivity workers also receive the value of what they produce, or $f(\tau^*) = \tau^*$. Thus, the presence of asymmetric information implicitly transfers resources from the more productive to the less productive in this model. The asymmetric information also reduces aggregate productivity (and thus aggregate output) because $\tau^* > \hat{\tau}$.

Finally, this analysis ignores perhaps the greatest beneficiaries of the firm-provided training: consumers. By lowering costs of production, the training results in lower costs of consumption arising from the pressure of competitive product markets.

5.3.2 A simple example

Let us keep the production function as $f(\tau) = \eta_i \times \tau$ and let $c(\tau) = \frac{\kappa}{2}\tau^2$. With this parameterization, the solution to the social planner's problem reduces to

$$\tau^* = \frac{1-p}{\kappa}, \tag{35}$$

while the asymmetric solution satisfies

$$\hat{\tau} = \frac{1}{\kappa}\left(\frac{(1-\lambda)(1-p)p}{p+\lambda(1-p)}\right). \tag{36}$$

If we take the ratio of the asymmetric information level of training to the socially efficient level we obtain

$$\frac{\hat{\tau}}{\tau^*} = \frac{p(1-\lambda)}{\lambda + p(1-\lambda)} < 1. \tag{37}$$

The ratio declines in λ, i.e. the higher the exogenous level of turnover for high-productivity workers, the lower the ratio of the equilibrium training level to the socially optimal level.

Finally, we may use the parameterization to examine the conditions that generate positive wage growth. In notation, we want to know when

$$w_1 = \frac{(1-\lambda)(1-p)\hat{\tau}p}{p+\lambda(1-p)} - \frac{\kappa}{2}\Big(\frac{1}{\kappa}\frac{(1-\lambda)(1-p)p}{p+\lambda(1-p)}\Big)^2 < \frac{\lambda(1-p)\hat{\tau}}{p+\lambda(1-p)} = v(\hat{\tau}). \tag{38}$$

After a bit of algebra, we want to know when

$$\Delta = v(\hat{\tau}) - w_1 = \frac{\hat{\tau}(1-p)}{p+\lambda(1-p)}\Big(\lambda - \frac{1}{2}(1-\lambda)p\Big) > 0. \tag{39}$$

Equation (39) implies positive wage growth when

$$\lambda > \frac{p}{2+p}, \tag{40}$$

so that if, say, $\lambda = 0.2$ then we need $p < 0.5$ as well. Large values of p imply low levels of training in the asymmetric information regime as well as in the social planners' solution. Of course, increases in λ also reduce the level of training in the asymmetric information regime. Realistically, at this level of abstraction, we cannot say much about the real-world plausibility of the condition in (40), but it is certainly easy to construct examples wherein the model predicts wage growth.

5.4 More empirics

5.4.1 Test for asymmetric information

AP (1998) provide a major test of their framework by utilizing a compelling institutional context to provide evidence of the substantive importance of asymmetric information in the labor market. The institutional context combines the German apprenticeship system with the German military draft. AP set the stage for their analysis by making the case that the apprenticeship system provides mainly training in general skills via institutional description and survey evidence on worker reports of how much of their apprenticeship training they use at the non-training firm. They also present evidence that firms, particularly larger firms, incur substantial training costs under the system, though much of the available evidence relates to the incidence of accounting costs rather than economic costs.

The reader may find it strange that AP use the German apprenticeship program to study firm training. After all, Becker's model applies to competitive labor markets, whereas the German apprenticeship system consists of a set of formal and highly regulated institutions.[82]

Yet one specific feature of the system renders it ideal for studying adverse selection. Apprenticeship contracts have fixed duration. When the contract ends, firms may, but need not, retain an apprentice by entering into a regular employment contract with them. This feature allows firms to shed less productive workers without having to dismiss them. These retention choices by firms result in adverse selection of less productive workers into the post-apprenticeship labor market.[83] Of course, workers who change firms also lose the benefits of their firm-specific training. To the extent that workers and firms share the returns to such training, an observed wage penalty for workers who leave the firm that provided their apprenticeship combines the loss of firm-specific skills with adverse selection, making definitive interpretation challenging.

AP rely on a second feature of the German economy to separate the adverse selection wheat from the loss-of-firm-specific-skills chaff: the draft. During the period AP study, Germany still used conscription, (more or less) randomly choosing a subset of young men to serve in the military or in alternative civilian roles.[84] AP take advantage of the random draft by comparing men who left their apprenticeship firm for military service to those who stay with their apprenticeship firm and, separately, to those who leave for other reasons, conditional on various observed characteristics.

Data from the German "Qualification and Career Survey" (QaC) conducted in

[82]The German and other central and northern European apprenticeship systems differ from apprenticeship in North America due to their substantially wider occupational and industrial coverage and due to the leading role of firms, operating at the industry level, in their design and operation.

[83]To the detriment of perfectly productive workers who decline retention offers with poorly managed firms or who seek to leave the firm they apprenticed with for reasons not related to their own productivity.

[84]Germany left Friedman's (1962) memo on the merits of a voluntary military sitting in its national inbox until July 1, 2011, when it suspended conscription.

1979, 1985–1986 and in 1991–1992 underlie the empirical analysis. Focusing only on men (eligible for the draft) at larger firms (which appear to pay larger training costs), they find that stayers and military quitters have higher wages than those who leave the apprenticeship firm for other reasons.[85] The results lack statistical precision, likely due to what econometrician Art Goldberger liked to call "micro-numerosity": the analysis sample includes only 209 military quitters and only 631 apprentices who leave for other reasons. Still AP provide reasonably compelling evidence of adverse selection in a context where firms appear to provide general training.

5.4.2 Does "poaching" of workers limit training?

Acemoglu (1997), and Acemoglu and Pischke (1998, 1999a) all argue that monopsony in the labor market may result in an under-provision of training in general skills. In one sense, this is undeniable. If firms have monopsony power then they hire fewer workers and so provide less general training *in total*, even if they provide the same amount *per worker*. The models in these papers also imply that firms offer less training *per worker*. As with our greasy spoon diner in Section 5.1, monopsony in the labor market provides an incentive (at the margin) for firms to hire experienced workers away from other firms rather than training new workers in a job unfamiliar to them (who then become targets for other firms looking for experienced workers who require less training). Monopsony implies that experienced workers at other firms receive wages equivalent to less than the value of their marginal product, so that the hiring firm can offer them a bit above their current wage and still make a profit.

Can this "poaching" story drive an equilibrium with less firm training per worker? Maybe. Acemoglu and Pischke (1999a) offer a model that implies a switching point between an equilibrium in which firms finance general training with no worker contribution and an equilibrium where workers pay for all general training. In contrast, in the Stevens (2001) model, which relies in part on imperfect capital markets, no such

[85] If the military provided useful training the return to this training would invalidate their test. Both the institutional setup (short service in dull tasks) and the ancillary evidence of negative returns to military service AP present lead us not to worry on this account.

switching takes place.[86] Indeed, the presence of frictions in both labor markets *and* capital markets drives the results in her model. Stevens shows that with sufficient imperfections in both capital and labor markets, an increase in the firms' monopsony power can increase training, resulting in a welfare increase for both firms and workers. But sufficiently small capital market imperfections imply that an increase in firms' monopsony power can reduce training.

Whether we have discrete switching as in AP (1999) or a smoother transition as in Stevens (2001), trying to document the nature of the "poaching" externality presents a nuanced problem for the researcher. Too little friction sends us toward the perfectly competitive market that Becker (1964) envisions wherein firms provide (and workers pay for) general training at efficient levels. Additional frictions may increase or decrease training per worker depending on the nature and extent of the frictions and, more broadly, on how exactly the labor market works. And measuring frictions poses its own challenges. Given these difficulties, the literature's slow progress in documenting the "poaching externality" does not surprise us.

Much of the empirical effort centers on how monopsony power affects the provision of apprenticeships; see Mohrenweiser (2016), Mohrenweiser et al. (2019), Muehlemann et al. (2013), and Stockinger and Zwick (2017). While surely of interest in their own right— apprenticeships contribute an important share of the overall stock of human capital, particularly in Europe—in our view they do not provide fertile ground for testing general theories of poaching. Firms, industries and governments typically structure apprenticeships so that workers pay (most or all of the) costs of their general training via lower initial wages, with the apprenticeship contract specifying when the apprenticeship ends and the firm must either promote or release the worker. This institutional structure severely complicates the interpretation of findings in relation to theories that presume a competitive labor market (with or without frictions).

Turning to studies that do not involve apprenticeships, Rzepka and Tamm (2016) compare the training practices of firms in relatively isolated areas compared to firms

[86] We have more to say about imperfect capital markets in section 5.7.

that operate in areas with large numbers of competitors, building on the idea that firms in isolated areas have greater monopsony power. They find that firms in areas relatively dense with competitors offer less training. Of course, areas with more competitors also likely feature more already-trained workers available for hire. Croce et al. (2017) add a further complication: knowledge spillovers. The urban economics literature provides some evidence that such spillovers increase with density, complicating the use of density as a proxy solely for monopsony power. As they note, one can tell conflicting stories regarding the likely effect of knowledge spillovers on firm training choices, with some stories implying a positive effect on training.[87]

Picchio and van Ours (2011) pursue what seems to us a more fruitful approach. They look at changes in employment protection in the Netherlands, motivated by the view that employment protection regulations increase monopsony power. Using variation in exposure over time within firms, they find that employment protection regulation (and thus monopsony power) decreases training. Though statistically different from zero, they estimate substantively small effects, implying a small poaching externality.[88] Mueller (2014) offers another reason for empirically small poaching effects: payback clauses (a.k.a. "training repayment agreement provisions") in employment contracts whereby workers agree to compensate firms for training received should their employment relationship end prior to reaching a specified duration. Further research on these provisions by economists would add value.

Finally, we should mention the provocative findings in Benson (2013). He considers hospital support of nursing schools and their faculty members. Clearly, nursing schools provide general occupational training that allows their graduates to work in hospitals, doctor's offices, nursing homes, schools, or anywhere else that hires nurses. In many

[87]Croce et al. (2017) use Italian data to study the effects of both worker density and the fraction of nearby employers with university degrees on the training choices of small and large firms. They find that both decrease training for small (less than 50 workers) firms but have no detectable effect for larger firms. We would have liked to see results that interacted the two variables of interest as well as a stronger substantive case for the use of lagged (by 20 years) values of these variables as instruments, a decision that assumes more persistence in them than in the unobserved factors leading to endogeneity concerns.

[88]Picchio and Van Ours also look at the consequences of anti-competitive regulation in the output market and find no evidence of an effect on firm training. Other work, such as Bassanini and Brunello (2011), finds that opening markets in Europe increased firm training.

communities, hospitals employ a substantial share of the available nurses, which should afford them some market power. Benson finds hospitals with more market power in this sense provide more support to local nursing schools and faculty. Of course, providing such support differs in important ways from directly purchasing or providing training for individual nursing students.[89]

5.4.3 Some new evidence from cross-country comparisons

Reading Acemoglu and Pischke (AP) (1998, 1999a, 1999b) makes it clear that they believe that U.S. firms provide less training to their workers than do firms in Germany and Japan. For instance, AP (1998, 80) argue that "The failure of the U.S. economy to generate as much training as Germany or Japan is sometimes blamed on the higher turnover in the United States (e.g., Blinder and Krueger [1996])." No data existed back in the late 1990s that could directly support such a claim. Instead, one had to make an indirect case that combined theory with related evidence. First, the literature at the time (and the literature now) provides little evidence that workers pay for general training with lower starting wages. Second, as AP note, Blinder and Krueger (1996) and others show that turnover rates in the U.S. exceed those in Germany and Japan. This difference likely results in part from stronger employment protection laws in Germany and Japan, and should, according to theory, imply lower training levels in the U.S. Finally, Germany and Japan had more compressed wage distributions than the U.S., and AP (1999b, Section 4.2) argue that such compression affords firms the opportunity to pay a portion of the costs of general training and reap a portion of the benefits.

The Programme for the International Assessment of Adult Competencies (PIAAC) data that we introduced in Section 4.4 provides a means to directly investigate differences in training levels across countries, and the correlations between training levels and features of national labor markets. Recall that the PIAAC public use files allow

[89] We wonder if hospitals use their support to gain more reliable information on their potential employees. One could potentially use data on the durations of nurses' initial employment spells to test this hypothesis.

us to compare responses from workers in 33 countries (including the U.S., Japan, and Germany) to very similar questions about training in the past 12 months.

We start our discussion of the empirical patterns in the PIAAC data by reminding the reader that we found in Figure 1 that U.S. firms provide both more formal training and more informal training than firms in Germany and Japan. Indeed, both Germany and Japan rank near the bottom of the "league table" for formal training. We find that puzzling, but the data speak clearly on this point.

We then correlate the PIAAC data on firm training levels by country with a measure of the strength of the employment protection regime in Figure 4, with separation rates (i.e. turnover) in Figure 5, and with a measure of income dispersion in Figure 6. The notes to the figures document our sources for these measures. The results will not please partisans of any of the models on offer in the literature. Figure 4 reveals a meaningful negative correlation between the strength of employment protection and firm-sponsored training ($\hat{\rho} = -0.29$). Figure 5 offers a strong positive correlation between firm training and separation rates ($\hat{\rho} = 0.61$). Finally, Figure 6 gives one result broadly in line with the models that AP propose by displaying a negative correlation between income dispersion and firm training ($\hat{\rho} = -0.22$).

Our analysis sticks to the shallow end of the pool. A deeper analysis would worry more about the difficulties with training measurement (which likely vary among countries for institutional and translational reasons) and would undertake multivariate analyses that could more plausibly isolate the relationships between training and specific aspects of national labor markets. That said, our simple analyses raise important questions about the ability of extant theory, whether of the Beckerian or the AP variety, to account for simple patterns in the cross-national data.

[INSERT FIGURE 4 HERE. Caption: Employment Regulation and Employer-Financed Formal Training Across Countries Note: Employment Protection Legislation (EPL) is the OECD's measure of strictness of employment protection - individual dismissals (regular contracts). The underlying scale runs from zero to six with higher scores representing stricter regulation (https://stats.oecd.org/Index.aspx?DataSetCode=EPL_R). EPL is measured in 2011. The participation measure comes from the answer to PIAAC question B_Q11 and is based on full-time, private-sector workers ages 25 to 59. Formal training is measured in 2011. The empirical correlation equals -0.29.]

[INSERT FIGURE 5 HERE. Caption: Separation Rates and Employer-Financed Formal Training Across Countries Note: The monthly aggregate separation rate comes from Table 2 in Hobijn and Şahin (2009). Separation rates are measured during 1990 to 2006. The participation measure comes from the answer to PIAAC question B_Q11 and is based on full-time, private-sector workers ages 25 to 59. Formal training is measured in 2011. The empirical correlation equals 0.61.]

[INSERT FIGURE 6 HERE. Caption: Income Compression and Employer-Financed Formal Training Across Countries Note: The income dispersion measure comes from Roser and Ortiz-Ospina (2013). Income dispersion is measured in 2011 except for Japan (2008), the US (2010), Israel (2012), and Mexico (2012). Consumption rather than income is used for Kazakhstan, Russia, and Turkey. The participation measure comes from the answer to PIAAC question B_Q11 and is based on full-time, private-sector workers ages 25 to 59. Formal training is measured in 2011. The empirical correlation equals -0.22.]

5.5 Lazear-Cavounidis-Lang models of training

In recent years, researchers have sought to relax the assumption that general human capital has but a single dimension. This effort parallels similar developments in the literature on formal schooling, which has seen greater attention to school quality (of which peer quality is one important component) as well as to school type, as with the now voluminous literature on college major choice and its implications for educational and labor market outcomes. In the firm training world, while some general skills, such as showing up on time and anger management, have broad applicability throughout the economy, others, though not firm-specific, have more limited applicability, as with the skills learned by the cook at our greasy spoon diner or the skills provided to software engineers at Google.

We consider two papers from this literature: Lazear (2009) and Cavounidis and Lang (2020). We begin with Lazear's paper entitled "Firm-Specific Human Capital: A Skill-Weights Approach." Lazear spins out the implications of a simple model with two skills, both general in the sense that all firms use them, and both specific in the sense that different firms have different weights on the two skills in their production functions.

Let (A, B) denote the level of investments in the two skills, and let the function $C(A, B)$ describe the cost of acquiring the skills, with $C_A(A, B)$, $C_B(A, B)$, $C_{AA}(A, B)$, $C_{BB}(A, B)$, and $C_{AB}(A, B) > 0$, where single subscripts indicate first (partial) derivatives and double subscripts denote second (partial) derivatives. A worker with skills (A, B) has output at firm i given by

$$\lambda_i A + (1 - \lambda_i)B = B + \lambda_i(A - B). \tag{41}$$

Each firm draws its value of λ_i, $0 \leq \lambda_j \leq 1$, from the distribution $f(\lambda)$, where λ_i determines the relative weights on the two skills in its production process.

Workers randomly match with a firm, call it firm zero with skill weight λ_0, at the start of the first period, prior to making their skill investment choice. The skill investment problem involves dynamic choice under uncertainty because the worker may switch firms in the second period. In particular, in the base version of the model, the worker receives a single offer from a random alternative firm j at the end of the first period. The worker also faces an exogenous probability q of losing their initial job at the end of the first period, in which case they necessarily take the alternative job, though with less bargaining power due to the lack of a competing offer.[90] The possibility of switching firms means that skill investment choices do not adapt to λ_0

[90]Lazear assumes that employment always dominates non-employment for his highly motivated workers.

as strongly as they would in a world with no potential for job switching in the second period.

Lazear determines wages in his model through Nash bargaining, although the same qualitative results follow from other standard mechanisms. Neither the worker nor the firm has private information, and so under Nash bargaining they agree to a contract that maximizes their joint gains. This implies only efficient (voluntary) turnover. In the case where the initial job implies optimal investments have $A > B$, turnover occurs when

$$(\lambda_j - \lambda_0)(A - B) > 0. \tag{42}$$

As a result, in the base case the probability that the match persists equals

$$F(\lambda_0)(1 - q), \tag{43}$$

where $F(\lambda_0)$ equals the probability that the incumbent firm's wage offer exceeds the alternative firm's wage offer while the $(1 - q)$ term reflects the probability of exogenous separation.

We highlight five predictions from the Lazear model. First, consider a modest extension that allows the number of job offers received at the start of the second period to vary across workers. Such variation enables the model to capture the notion of market thickness, i.e. the idea that the labor market offers workers in urban areas or in high demand occupations more options. Similar to the mechanism in Stevens' (1994) paper, as market thickness increases, the surplus associated with training increases, while profits received by the first period employer decline. In this variant of the Lazear model, the probability that the worker remains at the first period firm equals

$$F(\lambda_0)^N(1 - q), \tag{44}$$

where N denotes the number of second period offers. This term becomes small very quickly as the number of offers grows, meaning that the model has the empirically uncomfortable prediction of near-certain turnover in very thick markets.

Second, firms with relatively uncommon skill weights (i.e. with values of λ_j corresponding to regions with low $f(\lambda)$) should experience less turnover. Workers will adapt their skill investments to the first period employer's technology to some degree; doing so at a firm with unusual skill weights reduces the probability of obtaining a second period offer attractive enough to justify switching jobs. Third, and not too surprising given the second prediction, firms in the model implicitly pay for a portion of the cost of general training. Fourth, to the extent that detailed data on occupation and industry proxy for the skill weights λ and $(1 - \lambda)$, the model predicts that including occupation and industry in a wage equation will reduce the coefficient on tenure. A detailed inventory of the skills used on each worker's job, though more difficult to obtain, would reduce the coefficient on tenure even more. Firms also have an incentive to organize their production (and partially fund their worker's investment in) skills that are not widely used in their labor markets.

Finally, consider an extension wherein the exogenous job destruction probability q varies by firm. In that case, firms with a lower probability pay higher wages in the second period because the workers tailor their skill investment choices more closely to the firm's production technology than do workers at firms with higher probabilities. While not a particularly subtle prediction, it does raise an important broader issue: In the real world, workers hold much of their asset portfolio in human capital. Asking them to invest in additional training likely increases the riskiness of their portfolios. This makes risk-averse workers worse off, though of course this model and all the others we have reviewed assume risk-neutral workers.

Cavounidis and Lang (2020, CL henceforth) build on Lazear (2009) in their paper "Ben-Porath Meets Lazear: Lifetime Skills Investment and Occupation Choice with Multiple Skills." The title beautifully captures the focus of this paper. CL extend Lazear's model to allow an arbitrary number of skills, adopt a more general production technology and consider the dynamics of skill investments over the lifecycle, very much in the spirit of Ben-Porath (1967). As the (straightforward) extension to more than two skills complicates the notation without adding intuition, we focus on the two-skill case in what follows. CL posit that jobs come from the set C defined as

$$C = \{J : J_1^\sigma + J_2^\sigma \leq 1 \quad J_1 \geq 0,\ J_2 \geq 0\} \tag{45}$$

where $\sigma > 1$ and (J_1, J_2) define the skill weights associated with a specific job. For example, if $\sigma = 1$, the set of potential jobs consists of the right triangle defined by $\{(0,0),(1,0),(0,1)\}$, while if $\sigma = 2$, it consists of the northeast quadrant of the unit circle. In the limit, as σ gets large, the set of potential jobs becomes a square of area one.[91]

Workers come to the market with a vector of skills $S_i = (S_{1,i}, S_{2,i})$. CL (2020, 1409) "treat pre-market investment as exogenous [but] assume that the worker can arrive in the labor market with something other than the skills that are optimal for her." Workers produce

$$A_1 J_1 S_{1,i} + A_2 J_2 S_{2,i}, \tag{46}$$

where CL define the $\{A_j\}$ as "the productive efficiency of different skills, representing how the current technology uses each skill" and assume that $A_j >> 0$ for all j. In this enthusiastically partial equilibrium model, we interpret the $\{A_j\}$ as combining skill prices determined in the labor market with purely technological factors.

If workers choose their jobs to maximize the value of their production subject to the feasibility condition in (45), we obtain the Lagrangian

$$A_1 J_1 S_{1,i} + A_2 J_2 S_{2,i} + \lambda(1 - (J_1^\sigma + J_2^\sigma)). \tag{47}$$

The necessary conditions include

$$A_j S_j - \lambda \sigma J_j^{\sigma - 1} = 0 \qquad \text{for } j \in \{1, 2\}. \tag{48}$$

[91] In this instance $\sigma = 10$ produces an approximate square implying that $\infty \approx 10$ in this context.

Solving for the chosen values of $J_{j,i}$ yields

$$J_{j,i} = \frac{\left(A_j S_j\right)^{\frac{1}{\sigma-1}}}{\left(\left(A_1 S_{1,i}\right)^{\frac{\sigma}{\sigma-1}} + \left(A_2 S_{2,i}\right)^{\frac{\sigma}{\sigma-1}}\right)^{\frac{1}{\sigma}}}. \tag{49}$$

Putting these optimal choices into equation (46) provides the economic values of the workers' skills

$$V(S_i) = \left(\left(A_1 S_{1,i}\right)^{\frac{\sigma}{\sigma-1}} + \left(A_2 S_{2,i}\right)^{\frac{\sigma}{\sigma-1}}\right)^{\frac{\sigma-1}{\sigma}}. \tag{50}$$

At this point, we finally reach the Ben-Porath part of the model. CL allow their workers to make investments in human capital throughout their lives. In the simple two-period version of their model, they allow investments during the first period which increase skills in the second period by I_1 and I_2. These increases come at a cost $C(I_1, I_2) = I_1^\rho + I_2^\rho$, where $\rho > 1$.[92] If we let β capture discounting and ignore skill depreciation, workers select their skill investments to maximize

$$\beta V(S_1 + I_1, S_2 + I_2) - (I_1^\rho + I_2^\rho), \tag{51}$$

where $V(S_1+I_1, S_2+I_2)$ equals the worker's production in the second period given their skill investments (I_1, I_2) in the first period and the (possibly different) job chosen at the start of the second period. The second period job choice problem directly parallels the first period job choice problem already laid out, but with the skills augmented by the investments.

A few of the results follow immediately. First, it will not surprise the reader that in a version of the model wherein β_i varies across workers, those who value future earnings more (i.e. have a higher β_i) invest more than those who value future earnings less. Interpreting β_i as age, with lower values for older workers with shorter time horizons, the model gives the Ben-Porath prediction that investments decline over the life cycle. Second, investments in skills increase in the initial skill level, i.e. I_1 increases in A_1 and I_2 increases in A_2, despite the fact that investment costs do not vary with the

[92] They consider a continuous time lifecycle model of skill investment as well.

skill level. As CL (2020, 1412) explain, "A worker who has a high level of skill chooses a job that makes greater use of that skill. Knowing that she will be in a similar job next period, the worker chooses to invest more in the type of skill that she currently uses." In addition, the Cavounidis and Lang model incentivizes workers to over-invest in general skills in order to raise their wage at the incumbent firm.

We find the implications of two further model extensions worth noting: mobility costs and credit constraints. In the mobility cost extension, workers now face a cost m of changing jobs between the first and second periods. If we think of jobs as firms that make offers in the labor market, something we did not need to do in the base version of the model, then the first period (i.e. incumbent) firm will attempt to exploit the mobility cost by offering a second-period wage equal to the workers' best outside option minus m. Workers will anticipate this behavior and demand higher wages in the first period, but it still distorts behavior by altering workers' optimal skill investment choices. Given the dependence of the second-period wage on the outside option, workers may invest in skills of relatively little value at the incumbent firm so as to improve their outside option, even if they anticipate staying with the incumbent firm in the second period. CL extend the model still further to allow the incumbent firm to offer training to the workers and show that firms may want to do so in the presence of mobility costs. These strategic skill investments are not efficient. They arise solely due to the multi-dimensional nature of general training and represent a fundamental insight provided by the CL analysis.

CL also show that multi-dimensional skills add important new insights into the complex effects of credit market constraints on human capital investments. Consider an illustrative example inspired by their paper. Suppose we have a set of workers whose endowment features more of skill one than of skill two, i.e. $S_1 > S_2$. Suppose further that, given the available production technology, skill two produces so much more than skill one that in the absence of credit constraints all of these workers invest solely in skill two, going into debt to do so. They would then switch to jobs that reward skill two—we could think of this as switching occupations—in the second period.

Now add a constraint κ_i that caps the amount that workers can borrow, and assume for simplicity that it varies randomly across workers. For workers with a high enough κ_i, the constraint will not bind, and they will invest in the unconstrained optimal level of skill two. Workers with intermediate values of κ_i will still invest in skill two, but the credit constraint will oblige them to make less than the unconstrained optimum investment. Finally, workers with the lowest values of κ_i will invest only in skill one, because the credit constraint makes it impossible for them to invest enough in skill two to make it optimal to switch to a job that emphasizes skill two (i.e. to switch occupations). Their constrained optimal investment in skill one may not exhaust their available credit, and so they will appear as if their credit constraints do not bind, even though their credit constraint had a major effect on their investment choice and labor market outcome. The non-linearity in the response to credit constraints, plus the existence of workers who appear unconstrained while still suffering the effects of the constraints, represent further fundamental insights generated by CL's analysis of multi-dimensional general skills.

In sum, the models in Lazear (2009) and CL (2020) nicely demonstrate the extent to which the results from first-generation (or Beckerian) models of firm training depend on the assumption that general training has a single dimension. We look forward to additional theoretical developments along these lines.

5.6 The role of search and hiring costs

Much of the labor literature focuses on the costs of job loss to workers, yet firms lose when matches dissolve as well. Firms lose investments in specific skills and (still somewhat controversially) in general skills embodied in the departed worker. They also face substantial hiring costs for a replacement, costs that the search literature attends to but much of the rest of the labor literature does not. This section remarks on those hiring costs, and the ways in which they have changed as a substantial component of all job search activity has moved online. Both hiring costs themselves, and changes in their nature and distribution due to the internet, have implications for how we think about searches for experienced and/or trained workers as an alternative to hiring untrained workers and then training them.

In his review of imperfect competition in the labor market, Manning (2011) documents large costs to firms of hiring and training new workers, ranging up to over 11 percent of the expected wage bill. Economists worry that parties who pay fixed costs in relationships may find themselves subject to hold-up problems, as the other parties in the relationship attempt to exploit their investment. As we noted in Section 5.3, in the presence of substantial hiring costs, workers may attempt to hold up firms that pay hiring costs just as firms may attempt to hold up workers who invest in firm-specific skills. Which side gets the better of these dual hold-up problems in a given context will likely depend on the nature of the position and on broader conditions in the labor and product markets.[93]

The internet has fundamentally changed the recruiting process on both sides. The ability to submit an application to a firm electronically has greatly reduced the costs of worker search, especially for already employed workers. In an important paper, Faberman et al. (2022, FMST henceforth) summarize worker self-reports of search behavior drawn from surveys administered every October from 2013 to 2017.[94] Aggregating across the years, they find that fully 21 percent of employed workers reported looking for a job in the four weeks prior to the survey. In the same four-week period, unemployed workers report sending out an average of 8.5 applications while employed workers sent out an average of 1.1 applications.

Indeed, among those who report receiving employment offers in the four weeks prior to the survey, 40 percent were employed and actively looking for work, while an astounding 33 percent claim to have been employed and not actively seeking employ-

[93]Using Chinese data, Kuhn and Yu (2021) report relatively modest turnover costs in the retail industry, with about 25 percent of departures leading to a spell of short-staffing. Using German data on worker deaths, Jäger and Heining (2019) demonstrate substantial heterogeneity in the costs of this type of turnover. Of particular interest, they find that the deaths of workers in thinner labor markets lead to larger wage effects.

[94]They use an annual supplement to the Survey of Consumer Expectations fielded by the Federal Reserve Bank of New York. The population of interest consists of U.S. adults ages 18 to 64. They drop households with incomplete data on demographics or labor force status. Online supplements to their paper provide the survey questions, details of the solid-but-not-great response rates, and a discussion of “representativeness.”

ment.[95] FMST note that 2.5 percent of all employed workers report having received "unsolicited offers" in the four weeks prior to the survey, where unsolicited means not the result of worker-initiated contacts. Assuming all workers have an equal probability of an unsolicited offer in each period, this implies that about 28 percent of a firm's workforce would get an offer each year even if none actively looked for alternative employment. This is "poaching" with a vengeance.

We would expect the lower job application costs afforded by the internet to increase the number of applications firms receive. Our convenience sample of employers we have read about in the popular press or where we know someone involved with hiring indicates that employers feel overwhelmed by the onslaught of online applications. Kline et al. (2022) provide somewhat less impressionistic indirect evidence: In their resume study of the hiring practices of the 108 largest US employers, 83 of the 108 use an intermediary to run their online application system.

Three other issues related to the growth of internet-based job search warrant mention. First, we expect very heterogeneous impacts of internet search across firms and workers. While the large firms that Kline et al. (2022) studied mainly rely on internet search, smaller firms, such as independent retail stores and restaurants, may still find the costs of internet search prohibitive, and so continue to rely on more traditional methods. Second, we would expect the fact that firms differ in their search technology to lead to the sorting of workers by search technology as well. For instance, using data from the NLSY-97, Kuhn and Mansour (2014) document that workers with higher levels of education and who scored higher on the Armed Forces Qualification Test (AFQT) have higher probabilities of searching for jobs on the internet.

Third, the growth of internet search creates another massive measurement problem. Existing measures of the costs of search like those so nicely compiled in Manning (2003), surely became obsolete years ago. Internet search has a much different cost structure than the traditional search based on paper applications and in-person interviews. We

[95]The remaining offer recipients include the unemployed (about 15 percent) and those out of the labor force (about 13 percent). This does lead one to wonder about the substantive meaning of labor force participation in the current environment.

expect the rise of zoom and other remote work technology during the damn pandemic, along with the broader and at least in-part persistent move toward remote work that it engendered, to further change both worker and firm search behavior in important ways. How should researchers think about measuring the costs of search going forward? We look forward to reading new research on this topic.

Recall that in our discussion in Section 5.1 of whether firms choose the level of training or not, we mentioned that the decision to train or not largely comes down to the relative costs of training new workers versus locating workers with the relevant experience in the labor market. The revolutionary changes in job search over the last few years have altered the costs of locating such experienced workers. Whether this has translated into firms that use the internet for their recruitment activities hiring a more experienced workforce strikes us as an important and interesting empirical question.

5.7 The worker as a consumer: Training, consumption, and leisure

The models presented thus far provide many valuable insights while incorporating very simple models of the workers whose behavior they seek to describe: typically just income maximization in the presence of perfect capital markets. In this section, we offer a model of training choices for somewhat more complicated workers, while returning to the world where skills have but one dimension. The additional complication arises from the expansion of preferences to include both leisure and consumption and the expansion of choices to include both training and labor supply. We think this model allows us to better understand the economic reasons for worker decisions regarding investments in firm-provided training.

Consider a two-period model in which workers invest in general training during period one to increase their productivity in period two. Suppose that in period one, each worker's productivity equals $f(0)$, while in period two it equals $f(\tau) > f(0)$, where τ denotes the level of training. We assume concave payoffs to training, with $f'(\tau) > 0$

and $f''(\tau) < 0$, but convex training costs $c(\tau)$ so $c'(\tau) > 0$ and $c''(\tau) > 0$. Let (x_1, x_2) denote consumption goods that workers purchase in periods one and two at prices (p_1, p_2). Each worker supplies labor equal to $(1-\ell_1, 1-\ell_2)$, where $0 \leq \ell_t \leq 1$ indicates the proportion of the worker's time devoted to leisure in period t. Let $0 < \delta < 1$ be the workers' discount factor and let r denote the market rate of interest. If we let A indicate workers' initial assets and $R = (1+r)^{-1}$ the market discount rate, the budget constraint becomes

$$A + f(0)(1-\ell_1) + Rf(\tau)(1-\ell_2) - c(\tau) = p_1 x_1 + R p_2 x_2. \tag{52}$$

We assume separability of leisure and consumption in workers' utility functions, with $g(\cdot)$ the differentiable, strictly concave consumption component and $\phi(\cdot)$ the differentiable, strictly concave leisure component. This implies workers have value functions

$$V = g(x_1) + \phi(\ell_1) + \delta\Big(g(x_2) + \phi(\ell_2)\Big) \tag{53}$$

The first-order conditions at the optimum include

$$g'(x_1^*) - \lambda p_1 = 0 \tag{54}$$

$$\delta\ g'(x_2^*) - \lambda\ R\ p_2 = 0 \tag{55}$$

$$\phi'(\ell_1^*) - \lambda f(0) = 0 \tag{56}$$

$$\delta\ \phi'(\ell_2^*) - \lambda\ R\ f(\tau^*) = 0 \tag{57}$$

$$-c'(\tau^*) + R\ f'(\tau^*)(1-\ell_2^*) = 0 \tag{58}$$

In our model, both consumption and leisure are normal goods. This means that an increase in the initial assets of workers, A, leads to increases in both consumption and leisure. But according to (58), the increase in leisure in period two lowers the level of training, τ, because the increase in leisure lowers the payoff to training. In models that assume a fixed level of labor supply, like many of those discussed above, the initial level of assets has a relationship to the optimal training choice. A richer model with more time periods that allowed for an extensive labor supply margin would obtain a similar result. For example, workers who choose to retire earlier would also choose to invest less in training, all else equal. Similarly, in a version of the model with children, parents who choose to withdraw from the labor force temporarily to rear them would invest less in training than those who choose not too, all else equal.

A quick comparison of equations (54) and (55) shows that if the relative price of consumption remains constant ($p_1 = p_2$) and workers have the same discount factor as the financial market ($\delta = R$), then consumption remains constant over time ($x_1^* = x_2^*$). Of course, if workers differ in their discount factors, those with less patience than the financial markets ($\delta_i < R$) will chose to consume more in the first period ($x_1^* > x_2^*$), while those with more patience ($\delta_i > R$) will back-load their consumption ($x_1^* < x_2^*$).

The discount rate yields less obvious (perhaps even counter-intuitive) comparative statics in our simple model. Specifically, more patient workers (i.e. those with a higher δ_i) choose less training than less patient workers. More patient workers prefer relatively more consumption and relatively more leisure in the second period. Yet, equation (57) shows that more training implies more labor supply, and thus less leisure, in the second period. The comparative static result follows immediately.

Four additional points merit mention. First, models with fixed labor supply (e.g. models that assume full-time work in every period) necessarily miss this insight. Second, changes in the discount rate of the worker relative to that of the financial markets alter the relative price of consumption in the two periods. For instance, making the worker more patient by increasing δ_i lowers the cost of consumption in period two. Indeed, with perfect capital markets and with $p_1 = p_2$, the consumption of goods depends only on the ratio of the market's discount factor to the individual's discount factor. A similar line of reasoning applies to leisure, but with the complication that training increases the opportunity cost of leisure in the second period by making the worker more productive.

Third, perfect capital markets play a crucial role in these results. In period one, workers less patient than the financial markets will invest in training and spend down their initial assets (and possibly run up debts) in order to finance more first period consumption and leisure. This arrangement breaks down in the presence of imperfect capital markets for workers with initial assets limited enough that the capital market constraints bind. While the literature offers many ways to model imperfect capital markets, the most severe form simply shuts off the ability to borrow in the first period. In a world with no borrowing, our results completely reverse. Unable to borrow money to finance current consumption, impatient workers will only invest in training when it yields a high enough return to overcome their preference for current consumption.

Fourth, while we have not built uncertainty into the model, risk plays a pretty obvious role in training decisions. Before investing in job training, workers will already have made substantial investments in schooling. Tack on the investment in firm training and we see that human capital makes up a huge fraction of most workers' wealth. Moreover, workers typically have trouble diversifying their human capital investment risk. Returns to housing, often the second largest component of workers' wealth, generally have a strong positive correlation with returns to human capital, further increasing the riskiness of workers' portfolios.[96] Firm-specific training carries even more risk, as firms can and do fail, especially younger and smaller firms. Asking workers to assume a portion of the costs and returns to firm-specific training when management has much better information about the firms' prospects strikes us as a big ask. Our big picture point: treating workers as risk neutral agents likely misses an important component of the economics in the firm training investment problem.

5.7.1 An example

To develop some intuition, we explore a simple example wherein we assign specific functional forms to the objects in our model. To start, a worker's "full income" in the

[96] Davidoff (2006) documents that in areas with a stronger positive correlation, workers hold less of their wealth in housing.

sense of Becker (1965), denoted in period one currency, equals

$$F = A + f(0) - c(\tau) + Rf(\tau) = E_1 + RE_2, \tag{59}$$

where E_1 and E_2 represent expenditures out of full income on consumption and leisure in periods one and two, respectively. Assuming Cobb-Douglas preferences and letting β denote the weight on consumption in the utility function, the indirect utility function for period $j \in \{1, 2\}$ becomes

$$V(p_j, w_j, E_j) = \ln(E_j) - (1-\beta)\ln(p_j) - \beta \ln(w_j) \tag{60}$$

where $w_1 = f(0)$, and $w_2 = f(\tau)$. In the same notation, we have demand functions

$$\begin{aligned} x_1 &= \frac{(1-\beta)E_1}{p_1} \qquad & \ell_1 &= \frac{\beta E_1}{f(0)} = \frac{\beta E_1}{w_1} \\ x_2 &= \frac{(1-\beta)E_2}{p_2} \qquad & \ell_2 &= \frac{\beta E_2}{f(\tau)} = \frac{\beta E_2}{w_2} \end{aligned} \tag{61}$$

and labor supply functions $L_j^S = 1 - \ell_j$.

To reduce the number of first-order conditions, rearrange (59) as $E_1 = A + f(0) + Rf(\tau) - c(\tau) - RE_2$. Given the indirect utility function in (60) and the full income constraint in (59), we can write the consumer's problem as

$$\begin{aligned} V = \max_{(E_2,\tau)} \Big(& \ln\left(A + f(0) + Rf(\tau) - c(\tau) - RE_2\right) - (1-\beta)\ln(p_1) - \beta\ln\left(f(0)\right)\Big) \\ & + \delta\Big(\ln(E_2) - (1-\beta)\ln(p_2) - \beta\ln\left(f(\tau)\right)\Big) \end{aligned} \tag{62}$$

The first-order conditions for an interior optimum are:

$$\frac{\partial V}{\partial \tau} = \frac{1}{E_1}\Big(Rf'(\tau) - c'(\tau)\Big) - \delta\beta\frac{f'(\tau)}{f(\tau)} = 0 \tag{63}$$

$$\frac{\partial V}{\partial E_2} = \frac{R}{E_1} - \frac{\delta}{E_2} = 0 \tag{64}$$

A very little bit of algebra on (64) yields

$$\frac{E_2}{E_1} = \frac{\delta}{R}. \tag{65}$$

In other words, equation (65) shows that the ratio of expenditures in the two periods equals the ratio of the individual's discount factor and the financial market's discount factor. For those more patient than the financial market, expenditures increase over time, while they decrease over time for less patient workers.

Our main interest lies in equation (63), which has four features of interest. First, the final term in equation (63) reflects the labor supply effect of investing in human capital. Because human capital increases future wages, investment in human capital limits future consumption of leisure. If we fix labor supply, as in many of the models considered earlier in the chapter, then (63) becomes

$$\frac{\partial V}{\partial \tau} = \frac{1}{E_1}\Big(Rf'(\tau) - c'(\tau)\Big) = 0, \tag{66}$$

which embodies a more familiar marginal condition, and training increases relative to the case where workers choose their labor supply. Second, wealthier individuals also invest in less human capital via training because of the same labor supply effect. This result also disappears with fixed labor supply. Third, more patient workers experience the labor supply effect more strongly, because they care more about leisure in the second period. As a result, with flexible labor supply they train less, *ceteris parabis*, than less patient workers. Fourth, the predictions about the labor supply effect are extremely sensitive to the perfect capital market assumption, i.e. the assumption that workers can borrow all the money they want to finance training and first-period consumption at a fixed interest rate. We leave it to our clever and ambitious readers to develop new models of investments in training with imperfect capital markets.

Overall, this example nicely illustrates how human capital investment decisions depend fundamentally on the economic environment. An environment with perfect capital markets and flexible labor supply yield completely different behavioral predictions than an environment with no capital market (i.e. no borrowing or saving) and fixed labor supply. Finally, we observe that the level of abstraction inherent in our model removes any substantive distinction between firm-provided training and formal schooling. The cost of training, $c(\tau)$, in our model could represent payment to a firm via forgone wages or payment of tuition to a trade school. Thus, our caveats about the fundamental role of the economic environment also apply to decisions about investments in schooling. Whether analyzing firm-provided training or schooling, assumptions about labor supply flexibility and capital markets matter.

5.8 Training as treatment

5.8.1 Introduction

This section engages a genre of studies that treat individual instances of firm training as "treatments" using the methods of the program evaluation literature. Relative to less direct approaches, this strategy offers both advantages and disadvantages. The directness itself, linking changes in outcomes to particular measured instances of firm training, constitutes one advantage. The ability to rely on the now well-developed and often quite compelling applied econometric tools of the program evaluation literature represents another.

On the disadvantages side of the ledger sit a variety of measurement issues. For instance, all of the studies we review in this section build on survey measures of training spells. As we document in Section 3, such measures often have recall problems in practice, problems not always acknowledged by the authors of the studies that use them. Learning by doing does not really occur in spells at all, and so its effect necessarily lies outside this strand of literature. The many short spells of formal and informal training that the data show that workers experience, which may have a large cumulative effect on labor market outcomes, will typically not generate detectable effects when considered as individual treatments.

Although it applies program evaluation methods, the "training as treatment" literature has some disadvantages relative to the vast related literature on government-funded training for the unemployed. First, firms rarely if ever assign training at random or using discontinuities in observed running variables, leaving these typically compelling identification strategies off the table. And even if a firm did assign training in one of these ways, they probably would not tell outside researchers about it. The absence of random assignment also makes methods based on conditional independence assumptions less plausible in the firm training context, as it implies the absence of "within-study design" studies (e.g. Calónico and Smith, 2017) that learn about the required conditioning variables using experiments as benchmarks.

Finally, many of the first generation of "training as treatment" studies apply sampling rules and identification strategies that current researchers would find untenable. Some should probably have been left undone or done differently even by the standards of their own time. It is a testimony to the importance of the topic that the papers got written and got published (and, in may cases, accumulated respectable citation counts, too). In reviewing this literature, we aim to celebrate those who blazed the intellectual trail while still making sure that less careful readers do not come away from these papers believing that the treatment effect literature on firm training is somehow mature or settled.[97]

[INSERT TABLE 8 HERE]

[97] For reasons of space, and with heavy hearts, we omit the small literature that seeks to estimate the treatment effect of training on firm outcomes rather than on worker outcomes. The interested reader should consider papers such as Black and Lynch (1996), Almeida and Carneiro (2009), and Konings and Vanormelingen (2015). The Jones et al. (2012) paper considered, whose analysis of worker outcomes we describe in the preceding sub-section, also investigates some firm outcomes.

5.8.2 Older U.S. studies

One has to start somewhere. Lynch (1992) seems a good place to start.[98] She utilizes the NLSY-79 that (as shown in Table 1) contains the following questions: "In addition to your schooling, military and government-sponsored training programs, did you receive any other types of training for more than one month?" and "Which category best describes where you received this training?" Both questions were asked for up to three training spells during the previous year. During the years covered by her analysis, the NLSY survey only asks about training spells lasting at least four weeks. This restriction omits a large fraction of all firm training spells and, as a result, complicates comparisons between her findings and those in the remainder of the literature. In her analysis of the effects of training, Lynch distinguishes between training spells at the current employer and training spells at past employers, with the aim of learning something about the generality of training. She also separates completed and in-progress spells of on-the-job training at the current employer, with the aim of learning about whether workers trade lower wages for general training.

Within her analysis sample, 128 (4%) of respondents report participating in some kind of formal on-site job training and 450 (15%) report participating in off-site job training during the 80-83 sample window.[99]As detailed in the footnote, the two training indicators represent very coarse measures in the sense of McCall et al. (2016); put differently, they incorporate substantial "treatment heterogeneity." We highlight this issue because subsequent papers that study training using the NLSY-79 data follow Lynch's classification. So, for example, when "the literature" talks about off-site training it includes everything from associate's degrees to rehabilitation programs. On top of this, she has a third category with apprenticeship programs but the participation rate in such programs in the US and in her sample are both very low.[100]

[98]Starting with Lynch (1992) means relegating fine older studies such as Brown (1989b), Lillard and Tan (1986) and Mincer (1988) to this footnote. In addition, we do not cover every relevant paper published after Lynch (1992), but rather offer a curated selection of papers illustrating key issues in the literature.

[99]Google Scholar citations to Lynch (1992) now far outnumber its 578 (= 128 + 450) participants.

[100]On-site training comprises responses (8) company-provided training and seminars or (9) training programs at work not run by the employer. Off-site programs include responses (1) business colleges, (2) nursing

Lynch conducts her analysis using a somewhat restricted sample. The NLSY consisted of 12,686 respondents but she drops those still in school (obviously, given her question), the military over-sample, and college graduates. She further restricts her sample to respondents with wage data in both the 1980 and 1983 waves. The latter restriction allows her to examine changes in log wages between the two years. This reduces her sample to 3,064 individuals, all in their early twenties.

Lynch generates three sets of estimates. The first come from a parametric linear model of the log wage in 1983 estimated by OLS. The model includes the following conditioning variables: tenure (weeks), experience (weeks), years of schooling, the "local" unemployment rate, urban/rural, male, nonwhite, healthy or not, married or not, labor union member or not, number of jobs held since completing school, and industry and occupation indicators for the current job. Identification implicitly relies on a conditional independence assumption though the paper makes no explicit substantive case for said assumption. A modern reader might worry that some of the conditioning variables represent mediators whose values might be affected by training. The second set of estimates adds a Heckman (1979) "selection correction" term to the same linear model, which turns out to move the estimates very little.[101] The third set of estimates comes from the aforementioned first difference analysis, with the difference in log wages between the 1983 wave and the 1980 wave as the dependent variable and changes (where meaningful) in the same set of conditioning variables on the right-hand side.

The levels estimates reveal substantively large effects of previous off-site job training and of current on-site job training spells (of at least four weeks duration) on wages. The first-difference analysis, which does not distinguish between training at the current firm and at previous firms, yields no detectable effect of on-site job training but does

programs, (4) vocational-technical institutions, (5) barber and beauty schools, (6) flight school, (7) correspondence courses, (10) seminars or training programs outside of work, and (11) vocational rehabilitation centers.

[101] Identifcation comes from some combination of the bivariate normality assumption and dubious exclusion restrictions. As such, as most modern readers would, we assign these estimates zero weight. See e.g. the discussion and references in Bushway et al. (2007) for more on the normal selection model.

find a substantively meaningful and statistically significant effect of off-site job training on log wages of 0.002 per week. The average duration of the off-site job training equals about 41 weeks, so a spell of average length increases earnings by about 0.82, or about the same as a year of schooling.

Like Lynch (1992), Parent (1999) builds his study on the NLSY-79 training measures. Writing several years later, he incorporates survey waves up through 1991, thereby increasing the number of both persons (5,649) and person-years (over 29,000) available for analysis. Parent (1999) generally follows Lynch (1992) in defining his training measures, but with a couple of potentially important differences. He uses the same rule to aggregate the training types in the NLSY survey into on-the-job, off-the-job, and apprenticeship categories and, like Lynch, he separately aggregates weeks of training during the current job spell and from previous job spells. Of the 5,649 workers in Parent's analysis sample, 890 participate in on-the-job training and 721 participate in off-the-job training. The first coding difference relates to the handling of off-the-job training received while not employed, which Lynch includes as previous training but Parent omits. The second coding difference relates to a change in the NLSY survey instrument, which asked only about training spells at least four weeks in duration up to 1987 but asked about all training spells regardless of duration starting in 1988. Perhaps surprisingly, Parent simply uses all available spells, rather than, say, continuing to impose the four-week restriction in coding later survey years even though the data no longer require it or explicitly distinguishing between longer and shorter spells. Given the large number of short training spells in these data (and indeed in pretty much every data set) this decision matters.

Parent (1999) presents three sets of findings. The first set comes from log wage equations that include tenure and its square, and (real) experience and its square as well as a wealth of conditioning variables include demographics, Census region indicators and other context variables, and industry and occupation indicators. He estimates the wage equation using both OLS and instrumental variables, where the latter employs the Altonji and Shakotko (1987) scheme based on deviations from means described in more detail in Section 5.2.1. The IV and OLS estimates tell the same qualitative story for on-the-job training, namely that it increases earnings substantially, but not for off-the-job training, where only the IV estimates indicate a strong positive effect. Perhaps surprisingly, and in sharp contrast to what Lynch (1992) found, completed on-the-job training at the current employer has (statistically) the same effect per day on log wages as completed on-the-job training at previous employers. When Parent restricts his data to earlier years, the pattern in Lynch (1992) reappears.

The second set of findings seeks to shed some light on the enduring question of whether workers implicitly pay for training via lower starting wages. To look at this, Parent follows Lynch (1992) and compares workers with a training spell in progress at the first interview date associated with a given job spell to those without one, with the idea that labor market competition implies the incorporation of completed training spells in the wage, but not ongoing training spells. Using both indicators for ongoing training and the number of hours of training in the ongoing spell, Parent finds strong and statistically significant negative effects on log wages, where the wage refers to the actual starting wage for the 1986 wave and later and the wage at the time of the interview for earlier waves. In addition to the usual suspects, the model includes the total training received over the job spell (within the data) as a control for unobserved ability. The resulting estimates offer some modest evidence of workers bearing some of the cost of initial training via lower wages.

The third set of findings arises from a Cox proportional hazard model of the effect of training on job match dissolution. Parent argues that training that augments firm-specific skills may extend job matches whether or not it shows up in wages and finds some evidence in the data to support this claim.

Loewenstein and Spletzer (1999b) add value via their focus on direct measures of the generality of training. Workers provide an ordinal measure of training generality in the NLSY-79 while firms provide a measure of training generality in the EOPP.

Their NLSY sample comprises respondents to the 1993 and 1994 waves, employed but not self-employed, with valid values of the training variables and the outcome variables. They keep only the 91% of formal training spells paid for (according to the respondent) by the firm, and leave aside informal training due to measurement concerns. The 1993 NLSY survey asked respondents who reported a spell of formal training: "How many of the skills that you learned in this training program do you think could be useful in doing the same kind of work for an employer DIFFERENT than [current employer]?" A surprising 63% of individuals respond with "all or almost all the skills," the strongest of the five Likert scale options. Another 25 percent give the "more than half" and "half" responses. The EOPP sample retains all firms providing valid values of the variables required for the analysis. Employer respondents to the EOPP get asked: "How many of the skills learned by new employees in this job are useful outside of the company?" with possible responses of "almost all", "most", "some", and "almost none". A solid majority of 58% percent give the first response, while another 14 percent given the second. Loewenstein and Spletzer conclude "perhaps somewhat surprisingly in light of the emphasis that labor economists have placed on specific training, that the majority of employer-provided training is general."

In our view, really comprehending this evidence requires addressing three separate views of the specificity of training. The first, surely motivated by conceptual and rhetorical clarity, comes from the simple theoretical distinction (considered at length above) between general training, useful at every firm, and specific training, useful at only the current firm. The second comes from the empirical reality of the labor market, in which different jobs at different firms use different vectors of skills, with some of those skills general to all jobs (e.g. showing up on time), some common to all jobs in a firm, some common to particular occupations, some common to specific machinery or software, and so on. In this more empirical view, generality of training represents a continuous variable that in a loose sense integrates over the portability of the training received in a given spell to all possible other jobs. The third view inheres in the specific survey measures of generality deployed in this paper. The NLSY question purposefully limits the nature of the generality considered by restricting the respondent to the "same kind of work," a restriction that strikes us as likely to lead respondents to overstate generality more generally. The EOPP question comes closer to the empirical view just described, as it implicitly considers any jobs outside the current employing firm, while leaving the respondent to perform the loose integration over other jobs in their own special way. These data on generality pose a real challenge to the literature, and imply the need for effort on both the theoretical side and the measurement side to capture training generality away from the simply but extreme cases predominant in the literature.

Loewenstein and Spletzer (1999b) worry that reports of the generality of training by firms that provide it and workers that receive it may overstate the case if information frictions such as a lack of formal certification limit the ability of subsequent employers to detect, evaluate, and reward the training. To confront their worries, the authors first show using the EOPP data that employers reduce training duration for workers they report as having more months of "relevant" experience but not workers with more months of potential experience conditional on relevant experience, a pattern primarily present only for jobs requiring mainly general skills. The authors then estimate models of wage levels and wage growth that incorporate the explicit measures of training generality or, in the case of the NLSY, a measure of training type (company training, seminars at work, seminars outside work, school training, or other training). The estimates from these models support their conclusion that firms reward both the general training they provide and the specific training they provide, but perhaps not as much as they reward training provided on previous jobs. We would add, though, that despite the rich conditioning set allowed by the NLSY-79 data, the estimated effects of individual training spells seem really large to us, a concern the authors do not address. We suspect the large estimates flow in part from formal training spells acting as proxies for omitted informal training, along with a bit of residual selection on unobserved variables.

A final set of analyses considers the relationship between training and mobility. Estimates using the 1993 training generality measure lack power while estimates using training type in 1993 indicate that receipt of company training discourages job mobility. The authors view the small magnitude of this effect as further evidence of the generality of most firm training.

Frazis and Loewenstein (2005) undertake a classic (i.e. prior to the arrival of machine learning in economics) model selection exercise aimed at determining a preferred functional form for training hours in a log wage equation. They obtain similar findings from the EOPP and NLSY-79 data sets; we focus on the latter as more compelling. The NLSY-79 analysis uses training spells from the 1979 to 2000 waves, with hours of training at the current job as the primary training variable of interest.[102] The dependent variable is the log hourly wage, with some observations dropped due to implausible values. All of the partially linear models considered include a rich set of covariates including a cubic in tenure, interactions between the tenure variables and age at job start, AFDC receipt, years of schooling, union status, other demographics, an individual fixed effect, a job spell fixed effect, and various measures of non-firm training and training on earlier job spells.

The authors consider a long list of potential functional forms for the training variable, including an indicator for any training, a linear main effect, a combination of the indicator and the main effect, a quadratic, the log of hours of training (plus one to avoid taking the log of zero), the cube root, and a fourier series. The $\bar{R}^2$ serves as the metric for choosing among the different functional forms. They do not consider any interactions between hours of training and covariates. While the Fourier series fits the best, the authors value parsimony as well as fit, and so opt for the cube root, which fits best among the remaining choices, as their preferred specification. The relatively modest differences in fit struck us as a surprising feature of the findings: the $\bar{R}^2$ ranges from 0.2040 for the linear main effect (less than for the indicator on its own!) to 0.2050 for the cube root. That the results differed so little between men and women for this cohort when estimating the models separately that the authors only remark on it in a footnote also surprised us.

[102] As the reader will surely recall, NLSY-79 interviews prior to the 1988 wave do not collect information on training hours for spells less than four weeks in duration. The authors impute training hours for these observations. Because the early waves do not collect information on who paid for training for any spells, the authors confine themselves to longer spells of "company training" and shorter (imputed) spells because firms usually pay for both in the later waves.

Following the specification search, Frazis and Loewenstein estimate the "returns" to firm training. Despite the quite modest differences in fit among many of the specifications, they imply substantively large differences in estimated returns, with implied rates of return at the median positive value of 60 hours of training ranging from 12 percent to nearly 200 percent. With the NLSY data, the preferred cube root specification yields a value of 149 percent using the full sample and 159 percent in the sample that omits outlier values of training hours. Like us, the authors find these estimated returns implausibly large. They attempt to account for their size in a variety of ways, including estimating models allowing for person-specific linear time trends as well as person-specific intercepts, which makes some difference, and looking at the interrelation of training and promotion, which does not make much difference. In the end, they caution readers against generalizing their estimates of the effect of training on the trained to the currently untrained.

The (relative) lack of traction for this paper in the broader literature surprises us. First, we do not recall any other papers that take inspiration from this one by adopting a cube root of hours measure for training. Second, and more broadly, we do not recall any other papers doing meaningful specification searches. Most papers that go beyond a simple training indicator instead adopt a linear main effect in training hours, which Frazis and Loewenstein (2005, 474) note "always fits the data poorly."

5.8.3 Older U.K. studies

The extensive analysis of Blundell et al. (1996) forms our point of departure for the UK-based studies. They use the National Child Development Survey (NCDS) to study the determinants and effects of employer-provided work-related training. The NCDS comprises a representative sample of individuals born in the United Kingdom the week of March 3-9, 1958 and subject to an ongoing series of (widely spaced) interviews since that time. The authors make use of data from the fourth (1981) and fifth (1991) waves to study the effects of training received between the ages of 23 and 33 on outcomes at age 33. The UK Centre for Longitudinal Studies reports a 70.9 percent response rate for the fifth wave; the analysis does not appear to use any weights. When estimating the treatment effects of training, the authors drop respondents who report full-time school enrollment, self-employment, or no employment in 1991 and who report no employment in 1981, as well as respondents with missing data on job or training spell start dates, leaving them with 1,601 men and 1,180 women.

Table 1 gives the wording of the two training questions from the NCDS fifth wave survey. The questions implicitly ("[s]ince March 1981") cover the entire decade between the surveys. The first question asks about courses meant to lead to qualifications. 64% of their sample has received at least some training of this nature. The survey collects detailed information (e.g. start and end dates, who paid, intended qualification, etc.) for up to two courses, choosing those associated with the highest qualifications for respondents reporting more than two. The second question asks about courses that aim to develop job-related skills without leading to a formal qualification. The survey collects detailed data on up to three of these courses, this time choosing the most recent ones as required. 58% report at least one such spell.

The survey design leaves the researchers with many degrees of coding freedom. Blundell et al. partition the courses from both questions into six categories. The first five categories include only employer-provided training courses (EPTCs), which they divide into on-the-job (site or premises, as you like) with the current employer, on-the-job with a previous employer, off-the-job with the current employer, and off-the-job with a previous employer, and courses associated with the start of a new job (inferred from the job and training spell start dates) at any employer. A sixth, residual category includes non-employer-provided work-related training courses and courses for which detailed information was not collected, where the latter may include some employer-provided courses. The variable for each category consists of a count of the number of courses in the category the respondent reported receiving between 1981 and 1991. A separate set of three variables count the number of lower, middle, and higher vocational qualifications received over the same period.[103]

[103]We found the text a bit unclear about whether the counts include courses associated with a vocational

The available data and the available literature similarly leave the authors with many degrees of freedom in terms of specification and identification. Blundell et al. derive their ultimate estimating equation from three underlying wage equations, one for the respondent's first job, one for the job whose wage they observe in the 1981 wave and one for the job whose wage they observe in the 1991 wave. The authors merit praise for making their starting point explicit as well as the assumptions they require along the derivational path to their ultimate estimating equation in order to deal with concerns about selection into training based on either transitory outcome shocks or permanent unobserved differences (but not on the person-specific impacts of training - it is too early for that). The data allow them to motivate some of their assumptions via specification tests.

Their preferred specification has the first-difference of log hourly wages in 1981 and 1991 as the dependent variable, with the nine training type indicators, school qualifications and degrees prior to 1981, region, employer size, union membership, and occupation in both 1981 and 1991, and the log wage in 1981, as independent variables. They instrument for the 1981 wage, 1991 occupation, and 1991 employment status by including control functions that use family background measures, age seven ability test scores, characteristics of the first job, and pre-1981 training and post-school qualifications as instruments.[104]

The marquee estimates, presented in their Tables 5.6 and 5.7, reveal substantively large and statistically significant effects of EPTCs and of higher vocational qualifications for men. These results appear with or without the instruments. For women, similar patterns emerge for qualifications, while deploying the instruments largely does away with the estimated training effects obtained without them. Consistent with the publication date, the authors interpret their estimates as common effects. The estimates for training for men, and the estimates for higher vocational qualifications for both men and women, turn out substantively large, perhaps stumbling into the dark

qualification that the reader took but for which they did not receive the qualification.

[104] Those were the days! Though valid under the assumptions that justify their econometric model, face validity concerns would surely doom such instruments today.

forest of implausibility in some cases. For example, for men, off-the-job EPTCs at both the current and the form employer raise wages by around six percent, or nearly the effect of a "year" of formal schooling, despite their much shorter average length. Of course, even taking the estimated effects at face value, the marginal trainee may well experience a much lower return.

Blundell et al. complete their analysis by examining systematic variation in training effects based on training duration, training timing within the decade from 1981 to 1991, the coincidence of training and promotion, and formal school qualifications. We highlight the duration findings here as the most interesting. They show (only for men due to sample size concerns) a mostly monotonic relationship between estimated effects and duration, but also further highlight their astounding magnitude. For example, they estimate that off-the-job EPTCs at the current employer lasting *less than a week* raise wages by nearly seven percent. Please sign us up for this magic training! More seriously, we think these patterns signal underlying substantive issues not addressed in the analysis and perhaps not addressable in a satisfactory manner with the NCDS data. While the magnitudes of the estimates concern us, we assign real weight to the overall conclusion that EPTCs from both current and former employers often matter, especially when they lead to recognized qualifications.[105]

Booth and Bryan (2005) make use of the British Household Panel Survey (BHPS) from 1998, 1999, and 2000; the same waves as in their companion paper, Booth and Bryan (2007), discussed in Section 4.3.1. They select full-time, private sector employees aged 16-65 years "with valid information on our main variables," which gives them 7,167 observations spread over the three survey years and 3,333 workers. Specification tests lead them to pool men and women in their analysis.

[105] Arulampalam and Booth (2001) provide a parallel analyses based on the same two NCDS waves. They apply a somewhat different selection correction technology in the form of a negative binomial hurdle model, and different instruments. They do not disaggregate the training as much as Blundell et al. (1996), devoting themselves instead to estimating the effect of at least one course (of three days or more duration) and the effect of additional such courses. Oddly, they find that the first course has a huge effect on wage growth over the decade while additional courses have no effect (by which we mean a substantively small, negative coefficient estimate, not even close to statistically different from zero). The authors struggle to explain this pattern, as do we, unless we view it as a failed specification test.

Table 1 presents the basic BHPS survey question. The authors count only spells for which they have detailed information (the three "most important" ones in each wave) and only spells the respondent identifies as increasing or improving their skills on their current job. They aggregate training over the available waves (out of the three) for each respondent, separating for each person-wave observation training received at the then-current employer and at previous employers. Given their coding choices, only respondents present in at least two waves will have training with a previous employer. Viewed from another angle, recent job changers receive all the previous training considered in the study. The descriptive statistics reflect this dynamic selection, as about 33 percent of their person-year observations report at least one spell of training at the current employer compared to just over three percent reporting at least one spell at a previous employer.

Their choices regarding sample definition and coding have at least two important implications for the interpretation of their estimates. First, observed training at the current employer likely proxies for training at the current employer prior to the sample period (for respondents not changing jobs during the sample period) and for training at earlier employers (for respondents who change jobs, but not during the sample period). Second, measured training at previous employers represents a non-random subset of all such training, which one might expect to lead to a downward bias, but also proxies for training at the previous employer prior to the sample period, which one might expect to lead to an upward bias. The text barely hints at these issues, which we understand not as fatal, but as fundamental to interpreting their findings.[106]

The estimates arise from a fixed effects model of log hourly wages that includes some training measures along with a variety of conditioning variables, including quadratics in experience and tenure, local unemployment rate, and indicators for one-digit industry and occupation, region, marital status, firm size, contract type (permanent or tem-

[106]Unlike some other data sets, the wording of the training question in the BHPS implies that the authors' measure could include government funded vocational training (e.g. the "New Deal" in the UK at the time) leading to related employment. Note that the measurement issues differ substantively based on the number of waves included for a given respondent, though the paper neglects to pursue this aspect of the problem.

porary), union status, highest educational qualification and wave. Thus identification presumes conditional bias stability, a fancy way of saying that any remaining selection into training or into employment depends solely on time-invariant unobserved characteristics. The authors sometimes write as if they measure only conditional associations while at other times writing causally. We share their ambivalence.

Column (3) of their Table 1 measures training as counts of courses at the current and previous employers, and so parallels to some degree the specifications in Blundell et al. (1996). Booth and Bryan find that each current employer-financed course raises wages by about 2.4 percent while each previous employer-financed course raises them by about 7.8 percent. They interpret the difference as causal and thus relevant for choosing among alternative theories. Given the dynamic selection and proxy issues, we hesitate to pay much attention to the relative magnitudes but do see the estimates as supportive of the general view that firm training raises wages non-trivially for the trained. Additional estimates find that incidence does not over-shadow the counts (i.e. it is not all about the first course) for training at the current employer but does for training at the pervious employer, which feeds our proxy worries. In these data, and that, as in Blundell et at. (1996), the effects concentrate on courses leading to qualifications.

Dearden et al. (2006) add some novelty to the firm training literature by considering the effect of training on "direct" measures of productivity as well as on wages, both aggregated to the (roughly) three-digit industry level. The parallel examination of effects on productivity and wages allows their comparison, which potentially sheds light on models of firm training and the labor market. The construction of industry-level estimates and their comparison with individual-level estimates offers some evidence on the importance of externalities from firm training, something neither we nor the literature talk much about.

The analysis data from the UK Annual Census of Production and the UK Labour Force Survey (LFS). The Census of Production allows them to construct their "value-added per worker" productivity measure and also provides some conditioning variables.

The LFS provides the training measure as well as conditioning variables. They aggregate the individual-level data from the LFS up to the industry-year level. Because of data quality issues with the service-sector valued-added estimates, the authors confine their analysis to the production sector. Changes in classification codes and self-imposed cell-size minima lead to an analysis sample of 94 industries (or industry groups) for (at most) 14 years from 1983 to 1996.

The LFS asks:"Over the 4 weeks ending Sunday... have you taken part in any education or training connected with your job, or a job that you might be able to do in the future?" This question combines firm training with education and government-provided vocational training. It represents a flow rather than the stock called for by the authors' model. They repeat their analysis using both the flow measure and a stock obtained using the perpetual inventory method and obtain similar results. The follow-on question about training duration has a high non-response rate and other issues that lead the authors to avoid it, but they do note a median duration of two weeks. Over the period of their analysis, between eight and 14 percent of respondents report participating in job-related training.

Dearden et al. derive their estimating equation for the production function from a simple structural model, as traditional in that literature. The specification of the log-of-the-average wage equation then parallels that of the production function. They provide three sets of estimates, one from an (implausible) random effects specification, one from a within estimator that presumes only selection on time-invariant unobserved industry characteristics, and a posh GMM estimator that uses lagged levels as instruments for current differences to deal with training responses to transitory shocks. The authors clearly lay out the assumptions underlying each approach, and present some statistical tests supportive of the last one. In practice, the within and GMM estimators produce similar estimates.

The empirical work offers up two main findings: First, very large (probably too large, though the paper gamely tries to make the opposite case) effects of training on both productivity and wages. Second, substantively larger effects of training on productivity than on wages, which, as the authors note, could result from training externalities within industries, from cost-sharing for specific training, or from the sorts of labor market frictions discussed in Sections 5.3 and 5.4. The authors implicitly interpret their estimates as common effects, a stance consistent with their decision to weight the data by the LFS sample sizes available for each cell (and which implies many firms not spending enough on training).

5.8.4 Older German studies

To our knowledge, Pischke (2001) provides the first study of the treatment effect of training using the German Socioeconomic Panel (SOEP) data published in an English-language outlet. The SOEP is the German analogue of the U.S. PSID, and consists of a representative sample of about 4,500 households surveyed annually since 1984. The responses to questions from a continuous training module added to the 1989 wave of the SOEP provide the foundation for the analysis, which Pischke frames in part as motivation for the Acemoglu and Pischke papers we discussed in Sections 5.3 and 5.4.

As noted in Table 1, the opening training question from the 1989 SOEP wave reads: "There are various possibilities for work-related training. Thinking about the past three years, for your own job related education, have you read books or journals, participated in conferences and congresses, or participated in work-related courses?" For those reporting participation in work-related courses, the survey collects information about start date, duration (in categories), goals and whether the training took place during work hours or not for up to three courses. It collects more detailed information on content, costs, and perceived benefits about the single course the respondent identifies as most important. The SOEP posed its training questions to all respondents ages 16-64, no matter their labor market status, but Pischke restricts his analysis samples based on employment. This restriction has the effect of limiting the amount of publicly provided (via the UI system) and recreational training activities his measures capture. His Table 1 reveals that most training spells in the data last one week or less, though around 15 percent last three months or more.

Pischke (2001) estimates two different empirical models that build on two distinct identification strategies. The first model includes worker-specific fixed effects, motivated by an assumption that selection into training incidence and duration, conditional on a set of observed covariates, depends only on time-invariant unobserved variables. The second model includes worker-specific slopes in addition to the worker-specific fixed effects and presumes that any conditional selection on unobserved variables relates only to fixed unobserved differences or unobserved differences in factors that imply linear growth in the outcome variable. The second identification strategy nests the first but its additional flexibility requires an additional period of data (to allow for second differences) and consumes many additional degrees of freedom.[107] Both models condition on years of schooling, potential experience, tenure, an indicator for changing jobs and an indicator for full time work and both feature the log of earnings in the month prior to the interview as the dependent variable.

The analysis includes two training measures: one that captures incidence and one that captures duration. Both start at zero in the 1986 wave and cumulate over the waves from 1987 to 1989, with the fixed effects (and possibly the worker-specific trends) capturing the (effects of the) stock of training as of the 1986 wave. The incidence measure adds one in each year for which the respondent reports a training course. The duration measure sums the durations of the courses captured by the incidence measures, using the category midpoints to convert the categorical duration measure into a continuous one. The heavily skewed duration distribution, and the multiple categories for relatively short durations, temper the extent of the iatrogenic measurement error. On the other hand, Pischke's Table 3 shows that half of all respondents report receiving three or more courses over the three-year period, implying a specific sort of measurement error in both the training incidence and total training duration variables.

The empirical findings prove illuminating in the context of the broader literature. The fixed effect analysis yields imprecise estimates showing a substantively large earnings boost from training incidence of around 1.2 percent per course and a substantively

[107] See the related discussions in Heckman and Hotz (1989) and Moffitt (1991).

small earnings boost (relative to consensus estimates of the return to formal schooling) of 1.6 percent per year, both, alas, not precise enough to merit any stars. In stark contrast, including the worker-specific time trends, and thereby controlling for selection into training on (linear in logs) earnings growth moves the incidence effect close to zero while more the doubling the duration effect. Both estimates remain imprecise, and the duration effect remains low relative to the return to schooling, but Pischke (cautiously) interprets the changes in coefficients as signalling the importance of selection on earnings trends relative to the measurement error explanation for the puzzling incidence effects offered by Frazis and Loewenstein (2005).

Muehler et al. (2007) is the next German study published in English using the SOEP to study the effects of training. The 2004 SOEP survey contained the same training question as the 1989 wave. Using the responses to that question regarding training in 2001 to 2003, the authors construct three treatment variables capturing receipt of any firm (i.e. "continuous") training, receipt of any training in general skills, and receipt of any training in firm-specific skills. The latter distinction relies on responses to a question about the usefulness of the skills acquired in particular training spells at other firms, with the bottom two responses on a four-point Likert scale coded as specific training and the top two coded as general training. The authors highlight their separate estimates for general skills training and specific skills training as a contribution. The key interpretive challenge associated with their coding of their treatment variables arises from the fact that training incidence may arise from one, two, three, or more underlying training spells. Though not emphasized in the paper, this coding makes their estimates nearly impossible to compare to those in other studies. The sample includes only men, excluding usual suspects such as the self-employed as well as those with monthly earnings below €600 in the 2004 interview. The analysis sample of 1,751 includes 426 (24%) who report at least one training spell.

The authors present two sets of estimates based on conditional independence, emphasizing their ability to condition on prior training receipt as measured in the 2000 interview wave. One set comes from OLS estimation of parametric linear models, while the other from several variants of propensity score matching (with, oddly, a different set of conditioning variables). Both sets of conditioning variables include demographics and (of modest concern) variables like overtime work and job skill level *measured in 2004*. The authors preferred estimates come from difference-in-differences matching using log monthly wages from the 2000 wave as the pre-period outcome. The estimates from their preferred approach show substantively meaningful and statistically significant effects of training that arise almost entirely from general skills training.[108]

Görlitz (2011) published the first paper to study training using the German WeLL ("Berufliche Weiterbildung als Bestandteil Lebenslangen Lernens" auf Deutsch or "Further Training as a Part of Lifelong Learning" in English) data. As described in Bender et al. (2009), the first wave of WeLL data collection attempted surveys of 16,552 workers at 149 mid-sized (100 to 2,000 workers) establishments in the manufacturing and services sectors in Germany. In the end, 6,404 workers completed surveys from October 2007 to January 2008, for a response rate of about 39%. The author drops observations not employed at the time of the survey, with implausibly high or low values of monthly earnings, and with missing data on key covariates, leading to an analysis sample of 5,829. The paper does not mention the use of non-response weights. As displayed in Table 1, the WeLL survey asks about participation in training courses (including seminars) during the preceding two years. The author considers the effects of individual training courses (rather than the incidence of any course, as in the paper just considered) and does not distinguish courses by the specificity of the skills provided.

Inspired by Leuven and Oosterbeek (2008), Görlitz makes use of a special feature of the WeLL survey, namely its questions about intended courses. Specifically, the survey asks: "Did you intend to participate in training courses, seminars or lectures in the last two years without realizing this plan?" The survey collects data on up to three courses in addition to the question about unrealized but intended courses. Görlitz uses the responses to create eight categories: no training or intended training, only intended

[108] Ruhose et al. (2019) redo some of the previous German training analyses using the SOEP and expand the set of outcomes to include measures of civic/political participation, cultural participation, and social participation.

training, only one realized training, one intended and one realized training, and so on. She further restricts the unrealized trainings to those where the respondent offers a plausibly random reason for not fulfilling their training intentions, such as family or health issues. Overall, 63% of respondents report participating in at least one training course. Only modest numbers report unrealized training courses, but enough to make the exercise worth the attempt.

Estimation proceeds via a linear model with indicators for seven of the eight categories. Additional controls include sex, age, marital status, presence of children, years of schooling, tenure at the firm, indicators for white collar job, full-time job, and temp or fixed-term job, and firm fixed effects (all measured ex post). The unrealized courses act as the implicit comparison group for each level of training, including no training, where the author interprets the estimate as a measure of the selection effect, i.e. the extent of selection on the unobserved components of log monthly wages not removed by the included covariates. The firm fixed effects, enabled by the matched worker-firm study design, push the estimates toward zero. The data reveal a positive, substantively large, and statistically significant selection effect (i.e. non-trainees who intended training have higher earnings than non-trainees who did not intend training), and a lack of sufficient power to detect effects of meaningful size for the training courses. Görlitz defends, in solid casual Bayesian fashion, her estimate of 0.005 for the first course, a fair return for courses averaging only 38 hours of training, but one not statistically significant at conventional levels. The estimates for the second and third courses strike her as implausible.[109]

5.8.5 More recent studies

A handful of more recent studies illustrate the penetration of recent enthusiasms from the broader world of empirical economics into the "training as treatment" literature, including quantile regression, alternative outcomes, and, per the credibility revolution,

[109]For reasons of space, and because later waves of the WeLL featured even lower response rates, we confine ourselves to just one trip to the WeLL, other than noting that Tamm (2019) uses the WeLL data to look at the effect of training on the set of tasks workers perform on the job.

instrumental variables and randomized control trials (where randomization remains the best instrumental variable of them all).

Arulampalam et al. (2010) study heterogeneous effects of training by applying quantile regression methods to data on 10 European countries from the ECHP described in Section 3.2 and in our discussion of the Arulampalam et al. (2004a) paper in Section 3.4.1. By estimating linear conditional quantile functions at multiple percentiles of the wage distribution (10, 25, 50, 75, 90) in each country, the data can speak to whether or not training increases log wages more in some parts of the wage distribution than others. The study population comprises men ages 26-54 employed in the private sector in the 10 ECHP countries deemed by the authors to have sufficient, and sufficiently comparable, data. The resulting sample sizes range from 492 men in Belgium to 1,448 in France. The authors estimate their models separately by country but using the man-year as the unit of observation, where men persist in the sample for an average of about 3.5 waves.

The annual binary training measure equals one for man-years with a completed training spell during the recall window and zero otherwise.[110] In each man-year, the training variable included in the multivariate models equals the sum of the annual training measure up to that point in the data. The authors also include an estimated wage residual from the first wave to proxy for training (and other unboserved abilities and skills) obtained prior to the period covered by the data as well as an indicator for incomplete (due to censoring) training spells. This definition yields a remarkable spread in the incidence of training from 0.03 and 0.05 in Italy and the Netherlands at the lower end to 0.37 and 0.39 for Denmark and Britain at the upper end. This surprising (and to us rather concerning) variation occurs despite the harmonized ECHP survey instrument. The authors do not provide any information regarding the average duration (whether in days, weeks or hours) of these training spells due to concerns with missing start and end dates in some countries.

[110]The exact question wording appears in Table 1. Recall our discussion in Section 3 of the issues with the recall window in the ECHP.

The multivariate analyses feature the log of the hourly wage rate as their dependent variable and include a reasonable slate of conditioning variables: education, occupation, industry, age, tenure, recent unemployment, marital status, self-reported health status, type of work contract, firm size, year, and region. Estimating conditional mean functions, the authors find that average log wages of men with a spell of training in the preceding year exceed those of men without such a spell by about 1% in Denmark and Britain, by between 3% and 5% in Austria, Finland, France, and Spain, and by 9% in Ireland. They note that these differences negatively covary with training incidence at the country level. The authors do their very best to avoid causal language in describing these patterns, e.g. by talking about "ceteris parabis associations" instead of effects. We share their concerns about the plausibility of conditional independence given their covariate set, but suspect that they capture enough of selection to render measurement issues the primary concern at the margin.

Turning to their signature analysis, estimation of conditional quantile functions provides little evidence of heterogeneity in the "ceteris parabis associations" between training receipt and conditional quantiles of log wages. The combination of lack of power and measurement limitations imply that this null finding should not deter further search for heterogeneity in the treatment effects of training.

Brunello et al. (2012) apply instrumental variables methods to estimate the effect of what they term "continuous vocational training" in Italy. Their data comes from the Longitudinal Survey of Italian Households (ILFI, where the acronym refers to the name in Italian). The ILFI panel comprises five waves of data collection, once every two years from 1997 to 2005. The authors select individuals ages 20 to 55, employed in the private sector in a given year, and with valid earnings data as their population of interest. The sample consists of 1,874 workers with 4,719 person-year observations among them.

The ILFI questionnaire asks retrospective questions about training episodes in the two years prior to each interview. They define continuous vocational training as "any program organized by firms, local authorities and industrial associations that takes place after completion of upper secondary education and is not included in vocational tertiary education." Using the reported training episodes, the authors construct individual training stocks in each period using the "perpetual inventory" method.[111]

The instrument consists of conditional variation over region (13) and time (1999, 2001, 2003, and 2005) in the amount of public training grants from the European Social Fund (ESF) and the Italian government. The conditioning variables include age, sex, education, occupation, industry, part-time status, and firm size, along with region indicators and time-varying regional unemployment rate and GDP per capita variables. Thus, identification comes from conditional variation in the offered ("tendered" in the paper's terminology) training grants over time within regions. The authors point out that the unconditional variation does not follow a simple north-south pattern but omit an institutional case for exogeneity of the conditional variation.[112]

The first stage, which regresses the constructed individual training stocks on the covariates and the two-period lag of the training subsidy offers manages an F-statistic well above the usual cutoff for worrying about weak instrument issues. The estimates imply that one additional Euro of per capita spending on training grants in year $t - 2$ increases the stock of training in t by 1.03%. The OLS estimates imply that an additional week of training, in a model with a constant marginal effect of training, increases earnings by about 0.58% while the IV estimates imply an increase of 1.22%. Both of these estimates seem implausibly large to us when compared to, for example, credible estimates of the earnings effects of a year of schooling (roughly 30 weeks). The authors find that the effects concentrate on smaller firms and argue for their plausibility based on the view that they reflect highly productive marginal training at small, credit-constrained firms.

[111] The authors leave this method undefined in the paper. Some discussion of the issues associated with assuming a steady state in converting incomplete data on flows into stocks in the training context would add value to the literature.

[112] Indeed, the authors worry that the richer and better-run regions may, on average, attract more subsidies, but also have higher wages for other reasons. For instance, the Trentino region, one of the richest regions in all of Europe, managed to amass the most ESF training funds with €251.36 per capita, whereas the much poorer Puglia region in the south only managed to obtain €10.69 per capita. Perhaps the designers of the allocation system took their inspiration from Matthew 13:12: "For whosoever hath, to him shall be given, and he shall have more abundance" (KJV).

Jones et al. (2012) study training in a large group of co-op banks in Finland. They have rich administrative data on 223 banks for the years 2000-2004. The unit of analysis is the bank-year. The "Group Central" organization that unites the separately operated co-op banks provides subsidized general training to its member banks in Helsinki. This training, which the authors label "general", includes topics such as sales and the introduction of new financial products. Individual co-op banks also sometimes contract with outside consultants for training; the authors label this training "specific" and report that mainly upper-level managers receive it. Employees in these banks spend an average of 4.5 days per year in training, four days in general training and half a day in specific training. The paper does not provide any information on individual variation in training days. This training represents an accounting cost of around €1,000, split 60-40 between general and specific training.

The data for the multivariate analyses comprise branch-year observations. These analyses employ two measures of training: one the average training stock per employee in terms of expenditures and the other the average training stock per employee in duration (i.e. in days of training). As in the Jones et al. (2012) paper, the authors obtain stocks from flows using the "perpetual inventory" approach, again with no discussion of the potential issues associated with doing so.

Identification relies on an exogeneity condition that deviations from each bank's conditional mean training stock in a given year have a zero correlation with the unobserved component of outcomes. The authors state this assumption clearly but do not provide any institutional information regarding bank training decisions to support it. Implicitly, bank boards do the best they can, and some combination of local demand shifts, shifts in the board's beliefs about the production function and its tastes for training (the latter two perhaps caused by changes in board composition) generate the identifying variation.[113] The dependent variable consists of average wages in branch i in year t and the conditioning variables include the fraction of workers at the branch

[113] Estimates from a fancy dynamic panel setup do not differ much from the easier-to-interpret panel estimates we describe in the text.

with a secondary degree or more, mean worker tenure, mean worker age, fraction female workers, and the local unemployment rate.

The estimated elasticity of average wages with respect to the stock of training in expenditure terms equals 0.072, while measured in duration it equals 0.032. In both cases, entering "general" and "specific" training into the model separately indicates that the former generates most of the action. We hesitate to think of this as a causal effect of training (as the authors would prefer) because in an important sense, the bank boards jointly determine training and wages. Put differently, the economic meaning of the relationship between training and wages estimated in this paper strikes as unlike that in studies with many workers at different firms who change jobs more often than these low-turnover bank workers do.

De Grip and Sauermann (2012) implement a field experiment among workers in one department in a Dutch call center. That department handles calls from "private customers with fixed cell phone contracts" and employs 179 workers divided into 10 teams at the time of the study. In order to focus training expenditures on those thought most likely to benefit and least likely to leave the firm, management randomly assigned the most experienced workers to receive training either during the study period (the treatment group) or after it (the control group).[114] Due to limited training capacity, for many teams the treated units did not all receive training at the same time. A second dimension of randomization determined the timing of treatment, with many but not all teams having early and late training groups within their treated units. The authors use the period in between the training in the early and late groups to study spillovers from the early trainees to the late trainees.

Training in this study consisted of a one-week course on techniques to decrease the average time required to handle calls without lowering call quality. Even more than with the bank training in the preceding paper, the focus on a single department of a single firm implies a very homogeneous training treatment as well as presumably

[114]Due to illness or vacation, three treatment group members and nine control group member had their treatment status switched ex post. The analysis excludes these 12 workers, whose baseline characteristics do not differ statistically from those who remain.

error-free administrative data on training receipt.

Performance in this context equals the average time needed to handle customer calls. The sample mean of this measure equals 5.6 minutes; a decompositional exercise shows that worker fixed effects account for about 40% of the variation while period effects account for about 12%. The data include weekly performance for 32 weeks (covering 2008 and 2009): 17 weeks before training, five weeks during during which training took place, and 10 weeks after training. On average, workers appear in the data in 28 of the 32 weeks. The authors take the inverse of the average handing time in each week for each worker, multiply it by 100, take the natural log of it, and use the result as the dependent variable.[115]

Likely due to the small sample size, this study provides a rare example where including covariates in an experimental impact regression makes a real, substantive difference. The simple treatment-control difference in means equals 0.0122 while the estimates from the models with various conditioning sets equal 0.1127, 0.1244 and 0.0882. The authors summarize their findings as indicating a 10% productivity improvement from the training. Ancillary analyses suggest the effects decline with time and that spillovers from early trainees improve the productivity of later trainees before their training.[116]

5.8.6 What does it all mean?

In this sub-section we draw some general lessons from the "training as treatment" literature reviewed in the preceding sub-sections. As in the related summative exercise in Section 4.5, we focus on both what we do know and on what we know we do not know.

To begin, our take on the older literature. As we emphasized in Section 5.8.1, we

[115] Why not just estimate impacts in natural units? Two non-linear transformations seems like at least one too many to us.

[116] Sauermann and Stenberg (2020) use the less experienced workers (i.e. the ones not considered for training and so not randomized) from the De Grip and Sauermann (2012) paper as a non-experimental comparison group in a Lalonde (1986) style within-study comparison exercise. They find, as many such exercises do, that conditioning on pre-program outcomes substantially reduces the unconditional bias.

criticize the older literature not because we want to mimic adolescent historians who judge the people and institutions of the past by the fleeting standards of today and find them wanting. Instead, we ask a humbler and more useful question: how much weight should the older literature receive in thinking about training, and training policy, today. We answer that we would not put much substantive weight on the older literature for two main reasons. First, as noted in Section 5.8.1, standards for identification of causal effects have risen courtesy of the credibility revolution. Second, and arguably more important, technology and the labor market institutions have changed markedly in the developed world in recent decades. We noted some of these changes in our discussions of job search and of the production of training by firms. We think these pervasive changes imply only very modest temporal external validity for the older literature. Adding data quality as a third reason tempted us, but in fact, the quality of survey data on training has probably not improved much on net, with the general move to computer-aided interviewing improving data quality conditional on response, but with unit response rates falling, sometimes dramatically, over the intervening decades. An exception to the data quality point concerns studies that use administrative data, like the one about Finnish bank workers described in Section 5.8.5, though of course studies narrowly focused enough to rely on administrative data have generalizability issues of their own.

Caveats aside (just this once), in our view the existing empirical evidence (including the older literature) clearly demonstrates that training, general and specific, formal, informal, and casual, often matters for worker outcomes. We find this unsurprising, as training consumes real resources, which implies that both firms and workers have an incentive to invest in it only when they expect a reasonable payyoff. At the same time, we think the existing literature offers little in the way of credible evidence on the magnitudes of any causal effects of training, or about how those effects vary with characteristics of the training, of the firm, and of the trainee.

We made measurement a major theme of our chapter, and measurement issues bedevil the "training as treatment" literature in multiple ways. First, this literature, like the others we have reviewed, suffers from the measurement error in survey responses by both workers and firms that we document in Section 3. Second, this literature suffers from having not yet sorted out how best to measure the dosage of the training treatment, a task complicated by differences across surveys in what questions get asked. Should the training treatment consist of hours of training or spells of training or some combination of these, and in what functional form? Third, differences across surveys in what gets asked, in response rates, and in treatment coding, stymies the accumulation of knowledge across studies. Fourth, limitations of measurement imply that in most studies, the training whose impact the study estimates likely proxies for other training that either did not get measured in the underlying survey (such as casual training or learning-by-doing in our typology) or that the researcher neglects to include in the model, leading to biases of unknown magnitude. Some combination of administrative data and/or dedicated surveys focusing on training and related issues offer a way forward on this front.

Several substantive findings in (relatively) more recent studies stand out to us as worthy of further investigation. The first, uncovered by Pischke (2001) but also noted in a couple of other papers, concerns the potential for selection on trends (i.e. growth rates or trajectories) in outcomes rather than just on outcome levels. This finding has a parallel in the government-provided training literature as shown by Heckman and Hotz (1989). The second comes from the "old school" model selection exercise regarding the functional form for training hours provided in Frazis and Loewenstein (2005). They clearly reject a linear-in-hours-of-training model, but with little impact on researcher practice to date. Görlitz (2010) provides the third with her finding of a substantively meaningful conditional outcome difference between workers who intend to have training but fail to realize their intention and workers who did not intend to have training and successfully realize their intention. This last finding reminds us how much less we know about the appropriate conditioning sets for selection-on-observed-variables studies of firm training than we do in the adjacent context of government-provided training.

To date, the firm training literature has largely ignored treatment effect heterogeneity (a.k.a. "essential heterogeneity") in its investigations, other than sometimes producing separate estimates for men and women, or sometimes offering subgroup estimates for theory testing purposes as in Bellmann et al. (2017). The small sample sizes on offer in most of the available datasets rightly limit researcher enthusiasm for general subgroup exercises, whether traditional or operationalized via machine learning methods. At the same time, partly because of the age of many of the papers, much of the literature operate explicitly or implicitly in a common effect world, and so ignores distinctions between the average effect of training on the trained, the effect of training on the marginal trainee, the effect of training on trainees induced to train at the margin by a particular instrument (i.e. a local average treatment effect) and so on. These ideas, now painfully familiar in some literatures (e.g. government-provided training) but still sneaking up on others (e.g. public finance) should permeate the firm training literature going forward, due to their importance for correctly interpreting the empirical evidence. For instance, as noted in Frazis and Loewenstein (2005), a large estimate of the "effect" of training (i.e. something like the average effect of training on the trained) may or may not imply large effects of training at the margin, yet effects at the margin matter most for both firm and government decision-making.

While certainly related to the paucity of data available and the small sample sizes, we hope that nowhere else would applied researchers lump together such distinct treatments as flight school or business training with "inside seminars" and diversity training into a single catch-all training indicator. A small recent literature worries about what one might call the optimal coarseness of treatment measurement, i.e. how one should best chop up the dozens or hundreds of different training experiences a firm might offer into reasonable categories given the available data and a desire to obtain and present substantively meaningful estimates. McCall et al. (2016, Section 4.3) discuss the conceptual issues. VanderWeele and Hernan (2013) address this issue outside economics, while, more recently, Heiler and Knaus (2022) engage with the issue from an applied econometric standpoint.

Many papers in the "training as treatment" literature strike us as under-powered relative to training effects of the size one might expect. Though we recognize the tension between calling for finer measures of training in one paragraph and more statistical power in the next, we expect progress on both fronts. While the applied economics literature as a whole tends to do power calculations only for randomized control trials (RCTs) and not for non-experimental evaluations. In one sense this is natural, as the calculation can affect whether or not the researcher pays the fixed cost of data collection in the case of RCTs, while that cost is already sunk in non-experimental analyses of secondary data sets like the PSID or the SOEP. In those cases, the standard errors of the estimates implicitly provide an ex post power calculation. Still, in some non-experimental contexts, ex ante power analyses would allow researchers to avoid substantial data and analysis work that the data goddess would otherwise reveal as pointless ex post. Even the process of thinking systematically about a prior for "reasonable size" treatment effects in the empirical context at hand, as a power analysis forces the researcher to do, often adds substantive value.

The classic literature reviewed in sub-sections 5.1 to 5.7 emphasizes the conceptual distinction between fully general training, useful at every firm, and fully specific training, useful at only a single firm. In contrast, the "training as treatment" literature reviewed in this sub-section emphasizes typologies based on formality similar (or even identical) to the one we laid out in Section 2, relegating the generality versus specificity distinction to the background, or even omitting it entirely. While we appreciate the evidence on the relative determinants and effects of training categorized in these ways embodied in the extant literature, going forward we would like to see the literature incorporate skills useful at some firms but not all firms. An example close to our hearts may make our concerns plainer. Consider statistical software packages. Commonly used packages in our world include SAS, SPSS, Stata, and R. Knowing how to make effective use of one of these packages is neither a completely general skill, because most firms support only one or two, nor a firm-specific skill, as many firms use each of these packages. We also cannot call knowing how to use one of these packages an occupation-specific skill, as many occupations use each one, nor an industry-specific skill, as many industries use each one. Complicated? Yes, but we think many skills one might obtain via firm training have this character. Framed in terms of the existing literature, we would like to see ideas along the lines of those that that enliven Neal (1995), Loewenstein and Spletzer (1999b), Gathmann and Schönberg (2010), and other recent papers in the labor literature on skills seep over into the literature on the treatment effects of training via linkage with particular training episodes experienced by individual workers.

6 Policy

6.1 Rationales for Government Intervention

6.1.1 Motivation

Heckman et al. (1997) identify four potential reasons for government intervention related to human capital investment: (1) social externalities to poverty and unemployment, (2) credit market imperfections, (3) defective time preferences and shortsightedness of the poor, and (4) dysfunctional families that start children off with cultural and intellectual deficits.[117] For our purposes, and given our narrower interest in firm training, we combine (1), (3), and (4) into a single redistributive justification, while noting that the role of externalities in (1) and the role of normative judgements about the preferences of others in (3) make it a bit more than that. We also add a more recent (and very popular) rationale that builds on claims of monopsony power in the labor market leading to the potential "poaching" of trained workers, the threat of which then leads to under-provision of firm training.

6.1.2 Credit constraints

Consider (2) first. A large literature in economics addresses the empirical question of the extent of meaningful credit constraints in the market for human capital. Much

[117] They also identified a fifth: other market distortions, which is a bit vague for our list.

of this literature addresses one specific market: relatively recent high school completers considering enrollment in a two-year or four-year college. Key papers include Carneiro and Heckman (2002), Cameron and Taber (2004), and Belly and Lochner (2007). Lochner and Monge-Naranjo (2012) provide a survey. A smaller literature addresses credit constraints at later stages in the lifecycle and for workers mainly interested in less ambitious vocational training, as discussed in e.g. Baum et al. (2021).

An important feature of both literatures concerns margins. In most developed countries, both post-secondary academic training and adult vocational training receive heavy direct subsidies in the form of low (or even zero) tuition plus some mix of grants, loans, and guidance. All developed countries provide "free" primary and secondary school as well, where secondary schooling often has a direct vocational component. When considering policy proposals short of wholesale reformulation of the system, the relevant question concerns not the presence or absence of credit constraints in a world without all these existing programs, but rather in a world with them. Indeed, given these massive subsidies, one might legitimately worry that governments have subsidized human capital well past the point of efficiency.

As this brief summary hints, credit constraints matter mainly for training financed directly by workers, rather than provided by firms. Yet they remain relevant to our concerns here, even if not central to them. To the extent that firms under-provide training relative to the efficient level, whether due to monopsony and poaching or due to minimum wages making it impossible for low-skill workers to implicitly pay for training via lower wages, at the margin, the economy trades off between workers obtaining human capital at firms or via training outside of firms. To the extent that they obtain it outside of firms, credit constraints matter for us.

6.1.3 Poaching

This sub-section gives a policy spin to our formal discussion of the implications of labor market monopsony for training in Sections 5.3 and 5.4. As we described in that context, Stevens (1994) argues that the monopsony power of firms may result in under-provision of training. Numerous authors cite this "poaching externality" as a rationale for government intervention; see Katz et al. (2022), Brunello and Wruuck (2020), Hidalgo et al. (2014), Picchio and van Ours (2011), Görlitz (2010), and Leuven and Oosterbeek (2004) to name but a few. These papers often rely on poaching as a basis for advocating subsidizing firm training.

The intuition underlying poaching proceeds as follows: Monopsony in the labor market discourages workers from investing in training, as they anticipate the partial appropriation of the returns by their employer. On the flip side, monopsony power increases the incentive of the firm to invest in general training, precisely because their monopsony power allows them to reap some of the returns to general training. But a monopsonistic firm that invests in general training but does not pay the worker the full resulting increment in the value of their marginal product runs the risk that some other firm will hire the worker away by offering a higher wage. That threat in turn diminishes the firm's enthusiasm for providing the general training, with the result that the level provided falls below the efficient level that would emerge in a competitive labor market full of workers without credit constraints and without a minimum wage.

How does one design policy to circumvent the training shortfall resulting from monopsony? While "poaching" conjures up visions of illegal hunters killing rare, exotic animals, in this case poaching firms pay workers wages closer to the value of their marginal products. This presumably makes workers more likely to finance investments in their general human capital. Moreover, incumbent firms face poaching precisely because they pay their workers less than the value of their marginal products. Subsidizing the training of their workers is, in part, subsidizing the costs of those monopsonists. Is this really an economically justified way to expend taxpayers' dollars?

Even in situations where poaching appears empirically relevant, would a subsidy to training improve welfare? Moen and Rosén (2004) answer "no," at least within the confines of their model. They show that the level of training produced in a search model—and the resulting monopsony power—is optimal given the nature of the search environment, so that additional training would lower welfare.[118] Thus, for at least one model, subsidies to training would lower welfare in the economy. Of course, models like the one in Moen and Rosén (2004) have many moving parts. We would like to see the literature explore the related model space so as to determine the sensitivity of their welfare conclusions to specific details of their setup. In our view, this exploration should occur before we make any strong claims about the welfare effects of training subsidies to address poaching, whether positive or negative. Put differently, given the state of the (theoretical and empirical) literature as we write, it seems premature at best to argue that the "poaching externality" supports government subsidies.

Much recent research has been directed toward assessing the importance of monopsony in labor markets. Indeed, Card (2022), Berger et al. (2022), and Lamadon et al. (2022) all appeared in the *American Economics Review* in the first quarter of 2022, along with an entire issue of the *Journal of Human Resources* devoted to the topic. This literature faces daunting issues. Monopsony power may vary across occupations within firms. For instance, some local hospitals may have monopsony power in the market for nurses, but not in the markets for accountants or food service workers. Labor markets defined at the occupational level may lack relevance in thinking about monopsony power in some contexts. A dominant employer of entry level workers in the fast-food industry faces competition for such workers not only from other fast-food restaurants but also from employers offering entry level positions in many other industries. Moreover, as Card (2022) notes, the theoretical and institutional sources of monopsony power remain unclear.[119]

The appropriate levels of training subsidies and taxes must reflect these differences across occupations, firms and local labor markets. Put differently, a one-size-fits-all policy will likely fare poorly here in welfare terms. Designing a well-tailored policy of training subsidies strikes us as a major undertaking, fraught with many difficulties, including measurement of the monopsony power and degree of under-provision of train-

[118] Moen and Rosén (2004) do show that training subsidies in conjunction with other policies, such as taxes on "poaching" firms could, in fact, increase welfare in their model.

[119] Card (2022) and Manning (2021) offer recent reviews of this burgeoning literature.

ing for particular employers for particular occupations in particular labor markets, in addition to practical political and institutional concerns related to ensuring that the policy does not simply end up subsidizing training for politically favored firms. Given these difficulties as well as the challenges associated with measuring training at all that we discuss in Section 3, we are not very optimistic about economists' abilities to provide the necessary inputs for such a policy. Instead, we worry that this policy road leads directly to the land of wasted public resources and other unintended consequences. Standing back a bit, perhaps a better approach would seek to deal with any distortionary monopsony power directly, rather than focusing on the symptom of training under-provision.

6.1.4 Redistribution

In our view, we should justify government programs to encourage firm training mainly as a means of helping the less fortunate. The aid might take the form of, for example, financial help to allow disadvantaged workers to finance their training (as in current programs like the Workforce Innovation and Opportunity Act, WIOA, in the United States) or offering firms incentives to train (as in the sectoral programs considered below). Such programs, when designed well, embody tight targeting, unlike the indiscriminant subsidies implicit in policies such as "free" two-year college. Closer to home, we see no particular reason to subsidize human capital accumulation for economics professors no matter how monopsonistic their labor market. Given the focus on redistribution rather than efficiency, in evaluating such programs, we should simply ask whether they do a better job of improving the welfare of the disadvantaged than the equivalent in cash transfers would do (or at least a better job than an additional dollar spent on existing in-kind transfers such as public schools, childcare subsidies or SNAP).

6.2 Direct subsidies and tax relief

Unsurprisingly, there has been much experimentation by governments with different ways to encourage training, some good and some bad.[120] In our view, attempts to encourage training by offering broad "tax incentives" to firms to provide training, or by offering broad subsidies to firms that do provide training to their workers represent one notable experimental disaster. These broad training initiatives include the Dutch efforts in 1996 and 1997 studied by Leuven and Oosterbeek (2004) and the EU program described by Brunello and Wruuck (2020). According to Lynch (2000), several jurisdictions have experimented with training levies, which tax firms that do not spend "enough" on training, including France, Korea, Australia, Quebec, and Hungary.

Of course, to receive a subsidy or avoid a tax, firms must demonstrate that they actually offered training. As a result, tax incentives and subsidies should induce firms to alter the composition of training at the margin toward easier-to-monitor formal training and away from more-difficult-to-monitor casual or informal training. Black et al. (1999) find that larger firms (which provide more training per worker according to our review in Section 4) disproportionately rely on coworker training. As larger firms (and establishments) are more likely to have several workers performing the same task, they presumably find it cheaper for coworkers to provide the training rather than relying on more expensive managers (or external vendors). This line of reasoning suggests larger effects of government training incentives for smaller firms than larger firms, though at some level of subsidies, even large firms will formalize their training in order to avail themselves of the taxpayers' largesse.

But it gets worse. As documented in Section 4, firms provide a substantial amount of training to their workers even in the absence of any training taxes or subsidies. A general subsidy may increase training at the margin while at the same time paying firms for all of the training they would have provided anyway. The literature informally calls this "buying the base" or the "deadweight" associated with the subsidy. Subsidizing

[120] We use "experiment" here in the casual sense in which macroeconomists talk about "policy experiments" and cooks talk about "experimenting" with new ingredients; i.e. we use it to mean trying something new without doing any random assignment.

infra-marginal training seems to us an expensive means of encouraging more training at the margin, but subsidizing only marginal training presents such difficult administrative problems (similar in spirit to those associated with attempts to subsidize only marginal hires) that governments rarely attempt it. The following subsections summarize the available evidence on specific types of training tax and subsidy programs, including various voucher-based schemes.

6.2.1 Vouchers and grants

The state of Michigan operated the Michigan Job Opportunity Bank (MJOB) from 1986 to 1990. During this period, MJOB handed out over 400 grants with an average size of $16,000 (equivalent to $34,000 in 2022 dollars). The grants subsidized training in Michigan's manufacturing sector. Qualifying firms obtained grants on a "first-come, first-served" basis. To qualify, a manufacturing firm had to have fewer than 500 employees, not have received an MJOB grant in the past, and plan to implement "some type of new technology." The grants covered the direct costs of any formal or structured training program but did not cover foregone production or wages during training.

Holzer et al. (1993) make use of surveys sent out in 1990 to 498 firms that had all applied for the one-time training grants during 1988 and 1989. One in three (157) firms returned the survey, which contained information on total hours of training covering a three-year recall period (1987-89) for both successful and unsuccessful firms. Because the authors have all of the applications, they can investigate selective non-response on firm characteristics, and find very little. The grant-receiving firms report supplying, on average, 19 hours of training per worker per year in 1988 and 1989 compared to 11 hours among the firms that applied but did not receive a grant. This represents a deadweight loss of about 60 percent on the intensive margin. Disaggregating reveals that nearly the entire increase occurs during the year of actual grant receipt. For example, firms receiving grants in 1988 go from 7.7 hours of training per worker in 1987 to 36.0 hours per worker in 1988 to 10.1 hours in 1989. In other words, the grant induces a one-time boost in training. Consistent with the institutional setup and the general absence of selective survey non-response, the conditional estimates tell the same story as simple comparisons of means. The authors conclude their paper with a nice back-of-the-envelope calculation showing that each extra hour of training induced by the grants came at a government cost of \$6-\$7.

In 2006, the German state of Northern Rhine-Westphalia (NRW) introduced a training voucher program targeting firms with fewer than 250 employees. The voucher covered only work-related training provided by accredited training institutions. The voucher reduced the cost of the training by 50 percent per course (up to a maximum of €750).[121] Demand for the subsidy proved considerable, as the government issued 140,000 vouchers over the first 18 months of the program (in a population with somewhere between 1.8 and 2.5 million eligible workers). Following this initial surge, the state government tightened the eligibility requirements, which brought the monthly participation rate down to a third of its peak.

Görlitz (2010) studies the impact of the voucher program on training activities at the establishment level using data from the IAB establishment panel survey.[122] She estimates impacts on both the extensive margin (whether the establishment financed or provided any training during the year) and the intensive margin (the fraction of employees participating in any training) and considers three alternative identification strategies. The first identification strategy compares the before-after changes in these outcomes for eligible establishments in NRW and in three other German states (Lower Saxony, Hesse and Baden-Wuerttemberg) chosen to match the pre-trends in outcomes in NRW. The second identification strategy compares before-after changes for eligible NRW establishments with ineligible NRW establishments with at least 300 employees, where the gap from 250 to 300 acts to minimize slippage in eligibility status between periods. Both analyses apply traditional difference-in-differences (i.e. two-way fixed effect) estimators. The third strategy combines the first two and applies a standard triple

[121] An interesting paper would investigate the effect of the voucher program on course prices. The "Bennett hypothesis" in the education literature predicts they would increase to allow providers to capture more of the potentially available subsidy.

[122] IAB = Institut f ur Arbeitsmarkt und Berufsforschung, or the Institute for Labor Market and Vocational Research in English.

difference estimator. The analysis sample for the intensive margin outcome includes only establishments that provide or finance training, meaning that the two equations jointly represent a "two-part" model of sorts. Helpfully, the three identification strategies yield quite similar findings, namely a 4.4 to 5.9 percentage point increase in the fraction of establishments providing or financing any training and no detectable effect on the fraction of workers receiving training at such establishments. The lack of an impact on the intensive margin surprised us in light of the findings in Holzer et al. (1993).

In 2008, the German federal government instigated a similar program handing out vouchers that reduced training course prices by 50 percent, but with a lower per course maximum subsidy of €500.[123] Given the high demand for the NRW program, so high initially that it required a tightening of the eligibility rules, it might surprise the reader to learn about Görlitz and Tamm (2017), which describes an experimental evaluation of an information campaign to increase take-up of the federal voucher. The experiment randomly exposed eligible workers to information about the existence of the voucher, the conditions governing its use, and how to obtain it. The researchers implement a clever design that embedded the information shock within a broader telephone survey—"Employment today and tomorrow"—administered to both the treatment and control groups. They find that the information treatment increased knowledge of the program one year later by nine percentage points (or 30 percent) but had no discernible impact on take-up even though the experiment had the power to detect changes in take-up as small as 0.6 percentage points.

The literature documents two additional social experiments with training vouchers, one in Switzerland in 2004 and one a few years later in the Netherlands. The Swiss experiment, described in Schwerdt et al. (2012), selected a random sample of 2,437 respondents ages 20 to 60 from the 2005 Swiss Labor Force Survey (SLFS) as its treatment group. Treated respondents received a letter from the Swiss government

[123]The federal program differed from the NRW program in other ways, such as its prohibition on employers topping up the subsidy to cover the full training cost, and also changed over time.

in January 2006 containing an adult education voucher in one of three randomly selected denominations, equivalent to roughly \$200, \$750, or \$1,500. They could use the voucher for training courses, including leisure courses, starting by July 2006, but not for degree programs. In total, the treatment group members redeemed 449 vouchers, for a redemption rate of roughly 18 percent, with about 10 percent of redeemers using their voucher for a leisure course. Subsidized courses lasted an average of 42 hours.

The authors analyze the subset of those randomized who responded to both the 2006 and 2007 iterations of the SFLS. Due to normal rotation off of the five-year panel and due to a relatively high rate of panel attrition, the analysis sample contains only 1,422 treatment group members and 9,099 control group members. The treatment effect on self-financed adult education courses in the voucher year equals 12.9 percentage points on a control base of 33 percent, with some evidence that the voucher training crowds out firm training. Assuming the same redemption rate in the analysis sample as in the full experimental sample implies a deadweight loss of 30 percent. Surprisingly, take-up rates vary little with voucher value, with a rate of 0.10 for the least valuable voucher and rates of 0.14 the two higher values (though the authors lack the power to statistically distinguish the three rates). The authors do have enough power, however, to learn that compliers are much more likely to be highly educated; a result that should have been foreseen by the designers of the Swiss voucher scheme and which certainly speaks against the strategy of handing out vouchers at random from a redistributive perspective.[124] Finally, treatment effects on hourly earnings in 2007 and on employment in 2007 turn out very small and nowhere close to standard levels of statistical significance.

Hidalgo et al. (2014) report on the Dutch experiment. The study population consisted of workers in four sectors of the Dutch economy known to have a high percentage of low skill workers.[125] Within this population, the experiment randomly assigned 639

[124]We are reminded of the Epistle to the Romans 7: "I do not understand what I do. For what I want to do I do not, but what I hate I do." (NIV)

[125]The four sectors are: (1) Animal husbandry and greenhouse horticulture; (2) potatoes, vegetables and fruit; (3) food industry; and (4) natural stone.

of the 1,266 study participants to receive a €1,000 training voucher in 2006. The workers had two years to use the voucher. Assignment to the treatment group increased training at the extensive margin by almost 50 percent: a 20 percentage point increase relative to a control-group base of 45 percent. Presumably having two years rather than six months to use the voucher helps to account for the larger impact on take-up than in the Swiss experiment. Taking account of the (puzzling, and perhaps not generalizable to a permanent program) failure of some treatment group trainees to utilize their vouchers implies a substantial deadweight loss of 60 percent. The study finds no impact of assignment to the treatment group on log monthly earnings in 2008 (a coefficient of -0.004 with a standard error of 0.025) or on job mobility, defined as moving to another company (a coefficient of 0.01 with a standard error of 0.02).

6.2.2 Train or Tax

In 1995, the Canadian province of Quebec introduced the "1% law"[126] which required firms with annual payrolls above $250,000 to either spend (at least) one percent of their payroll on training or pay a penalty-tax of the same size to the Quebec Minister of Revenue. Only expenditures on general training, defined as "transferable skill-related structured training that is directly related to the job or that is recognized by other workplaces", counted under the law. In 2004, the law changed to exempt workplaces with payrolls below $1 million from the training levy. The exempted firms had lower compliance rates prior to the reform, with compliance defined as spending the required amount on training rather than paying the tax. For example, over the 2000-2003 period compliance rates for firms with payrolls exceeding $1 million varied between 86.8 and 88.2 percent, while those for firms with payrolls between $250,000 and $500,000 varied between 67.9 percent and 70.0 percent.

To investigate the impact of the exemption, Dostie (2015) uses the Canadian Workplace and Employee Survey (WES). As described in our Section 4.2.3, WES provides linked longitudinal data on employers and employees. The employer part of the WES

[126] "Loi favorisant le développement de la formation de la main-d'oeuvre" au francaise.

collects information about the number of workers receiving classroom and on-the-job training as well as payroll data. The WES notion of classroom training corresponds roughly to formal and informal training as defined in Section 2, while its notion of on-the-job training corresponds roughly to what we call casual training.

Dostie (2015) produces estimates based on three different identification strategies: (1) a simple before-after comparison of firms with annual payrolls between \$250,000 and \$1,000,000; (2) a difference-in-differences analysis using firms with payrolls less than \$250,000 and greater than \$1,000,000 as the comparison group; and (3) a triple-difference analysis that adds a comparison with Ontario to (2). Helpfully, all three strategies tell the same qualitative story and roughly the same quantitative story as well. Focusing on the triple-difference estimates, exemption from the 1% law reduced the average proportion of employees receiving classroom training by 0.071 and increased the average proportion receiving on-the-job training by 0.109, both statistically different from zero (but likely not from each other) at traditional levels. That WES contains data on establishments while the law applies to firms likely implies some attenuation bias due to misclassification. These findings strongly suggest that the Quebec training tax results mainly in the inefficient formalization of training that would otherwise happen informally, rather than actually increasing levels of firm training.

6.2.3 Vouchers combined with wage subsidies

Starting in 2006, the U.K. operated a very expensive program called Train-to-Gain that provided free-of-charge training in basic skills and vocational qualifications to employed workers with no or very low formal qualifications. In 2010, its year of termination, the program consumed about £1 billion, or more than a third of the government budget for adult skills and further education. In addition to covering the course fees (which ranged from £500 to £1,200), the program compensated employers for the wages workers received during time spent training, with the amount of the subsidy varying by workplace size and region of the country, and with a maximum total wage subsidy per worker of £525.

Prior to rolling out Train-to-Gain, the U.K. government operated Employer Training Pilots (ETP) in a small number of local areas between 2002 and 2006. The program implicit in the ETP differed in some small ways from the later Train-to-Gain program but not enough to raises concerns about generalizing the results from one to the other. Abramovsky et al. (2011) present the ETP evaluation commissioned by the U.K. government for an academic audience.

The ETP evaluation employs a difference-in-differences design using two sets of comparison areas chosen to match the pilot areas on particular dimensions. It makes use of three separate surveys, the first a special survey of eligible workers in two pilot areas and two comparison areas, the second a special survey of employers that employ eligible workers in eight pilot areas and the same two comparison areas, and the third the regular U.K. Labour Force Survey (LFS). The evaluators selected the comparison areas for the ETP-specific surveys to match the surveyed pilot areas in industrial structure and labor market performance. The broader geographic coverage and many years of pre-ETP data available in the LFS allow a different choice of comparison areas based on pre-trends in training levels and geographic proximity. The training questions on the ETP-specific surveys aim to pick up only that training subsidized by the program, while the LFS has more standard training questions that ask about a broader set of formal training experiences.

The evaluation focuses mainly on the effect of the ETP subsidies on training, rather than on later labor market outcomes. All three surveys, as well as a variety of sensitivity analyses including varying the set of the comparison areas in the LFS analysis, tell the same story of substantively small positive effects that never statistically distinguish themselves from zero. Even taking the larger of the point estimates at face value, the vast majority of the training subsidized by the program consists of deadweight, i.e. of training that would have occurred even in the absence of the subsidies. Given the findings, we wonder why the government went ahead with the Train-to-Gain program.

Finally, two studies have investigated the German WeGebAU (a play on the German word for road construction)[127] program, which began in 2007: Dauth and Toomet (2016) and Dauth (2020). WeGebAU pays for occupational retraining programs for low-skilled employed workers and for older workers (age $\geq$ 45) at smaller firms ($<$ 250 employees). The 176 local employment offices in Germany operate the program, which aims to improve skill levels among employed workers and to increase labor force participation among older workers. Under the program, the employer releases the employee from work to participate in general training but continues to pay their wages during the training period. The program covers the direct cost of the training along with reimbursement of related expenditures such as childcare and transport. For low-skilled workers, the program also provides wage subsidies to the firm. Any of the employer, the employee, or a caseworker can initiate an application for support, all must agree for training to proceed. The training subsidized by the program ranges from very specific and relatively short programs aimed at specific certifiable skills such as operating a forklift, to occupational training in health care or IT lasting many months. Surprisingly, only about 60 percent of firms participating in WeGebAU claim either cost reimbursements or wage subsidies, despite payouts as high as €4,000 for cost reimbursements and €6,900 for wage subsidies. Between 20,000 and 40,000 German workers participate in the programs each year but despite their generosity, the program never fully exhausts its annual budget of around €200 million.

Dauth and Toomet (2016) evaluate the older workers part of WeGebAU. They use register data on workers entering the program between July 2007 and December 2008 along with a dynamic treatment effects (DTE)[128] approach that assumes no anticipation effect and conditional independence. Their DTE approach estimates the impact of treatment this month versus no treatment this month but maybe treatment later. Given the very low take-up rate of around 0.2 percent over their 18-month interval, we (and they) expect that the DTE estimand approximates treatment versus no treatment in their context. They employ two comparison groups, one consisting of eligible

[127] The program later changed its name from "Weiterbildung Geringqualifizierter und beschäftigter älterer Arbeitnehmer in Unternehmen" to "Weiterbildungsförderung Beschäftigter."

[128] On DTE, see e.g. Sianesi (2004) or Crépon et al. (2009).

co-workers at the same firms as the treated workers and the other of eligible workers in firms that have not yet participated in WeGebAU. Their preferred estimates reveal that older workers receiving training via WeGebAU have a 0.03 higher probability of employment 18 months after training start.

Dauth (2020) evaluates the low-skill worker component of WeGebAU. In the spirit of Frölich and Lechner (2010), identification comes from residual differences across local labor offices in the number of workers who receive training subsidized by the program. The author refers to these residual differences as program styles, and thinks about them the way the economics of crime literature thinks about variation across judges in the severity of sentencing. This source of variation identifies a local average treatment effect (LATE), i.e., an average effect on workers who receive subsidized training from a generous local labor office but would not from a stingy one. For low-skill WeGebAU participants from 2007 to 2010, the analysis finds that over the five years following training start, participants experience 93 days of additional employment and an impressive 12-13 percent increase in accumulated earnings. Perhaps surprisingly, the LATEs differ only modestly from the corresponding OLS estimates. A back-of-the-envelope cost-benefit analysis shows higher costs than benefits, but leaves room for a broader analysis to come to a different conclusion.

6.3 Minimum wage and training

In the classic Beckerian model discussed in Section 5.1, workers finance their training by accepting a lower starting wage combined with a higher wage after training. Rosen (1972a) points out that a binding minimum wage eliminates this possibility, which may interfere with the ability of some low-skill workers to finance their general training.

Firms, however, have many margins on which to adjust job quality in response to a minimum wage increase, margins that may dominate cuts in training in many cases. Consider a restaurant that suddenly faces a minimum wage law that creates rents for their employees. It may convert full-time positions into part-time positions and fixed schedules into uncertain ones. It may reduce its crew size and push the remaining workers to work harder. It may cut free or reduced-price employee meals. It may invest in labor-saving technologies such as self-service drink stations or electronic order systems. Or it may hire workers with sufficient experience to require less training, a result in line with the standard (but surely less creative) model.[129] Which of these adjustments an employer chooses in a particular context will depend on the vagaries of the firms' technologies, the labor market conditions it faces, and the needs and wants of its customers. Clemens (2021) reviews the broader literature on how firms respond to increases in minimum wages on non-employment margins. We note the literature on other response margins because it informs our prior, which centers on rather modest effects of minimum wages on training.

Hashimoto (1982) authored the first published paper to examine the impact of the minimum wage on the the wage growth of lower-skilled workers. Using data from the National Longitudinal Survey (NLS) of Young Men, Hashimoto estimates substantially lower wage growth for workers with starting wages close to the minimum wage. As the NLS lacks data on training, Hashimoto uses his theoretical model to infer the extent of unmeasured training from the observed wage growth. This strategy generates an estimated reduction in training of between 2.7 and 15 percent. As Hashimoto notes, around the same time Fleisher (1981), Mincer and Leighton (1980), and Lazear and Miller (1981) all use the NLS data to look at the effects of the minimum wage. Using differing methodologies, Fleisher (1981), and Leighton and Mincer (1980) find that the minimum wage retards training, while Lazear and Miller (1981) obtain inconclusive results.

The NLS-based studies all had to rely on wage growth as a proxy for job training. Using the EOPP data discussed above, Grossberg and Sicilian (1999) provide the first study utilizing a direct measure of training. Consistent with the NLS-based studies, Grossberg and Sicilian do indeed find that workers paid the minimum wage have lower levels of wage growth than other workers and that workers paid the minimum wage

[129]We end our list of margins here but, to quote the Bard, "There are more things in heaven and Earth, Horatio, / Than are dreamt of in your philosophy." (Hamlet [1.5.167-8])

receive less training than the workers with wages just above the minimum wage. Yet when they look at workers paid less than the minimum wage, and hence not constrained by it, they find these these workers do not receive substantially different amounts of training than the workers at the minimum.[130] Grossberg and Sicilian interpret their results as a cautionary tale regarding the use of wage growth to infer training receipt and as consistent with the mixed findings in Lazear and Miller (1981).

Neumark and Wascher (2001) use the 1983 and 1991 training supplements to the Current Population Survey to estimate the impact of state minimum wage laws and the 1990 increase in the federal minimum wage. They construct their training measures from the CPS questions listed in Table 1 and the related follow-up questions. To capture some of the dynamics in a period when many states changed their minimums, they use as their primary minimum wage measure "the percent by which the state minimum exceeded the federal minimum over the previous 3 years." Their cross-sectional analysis uses workers ages 20-24 in the 1991 data as the treatment group and workers ages 35-54 in the 1991 data as the comparison group. The panel analysis uses workers ages 20-24 from the 1983 data as the comparison group. Both analyses indicate a substantial negative effect of higher minimum wages on receipt of formal training. Estimated effects on informal training turn out substantively and statistically insignificant in both cases. To our knowledge, they present the first estimates of the effects of the minimum wage on training obtained using panel methods (i.e. two-way fixed effects), implying identification via conditional bias stability.

Acemoglu and Pischke (2003) use the training data from the 1987 to 1992 survey years of the NLSY-79 to examine the effect of minimum wages on training for

[130]In the U.S., the federal minimum wage was set by the Fair Labor Standards Act (FLSA) in 1938, which had substantial revisions in 1961, 1966, and 1978. Prior to 1978, agricultural workers had a lower minimum wage or were uncovered. See U.S. Department of Labor, Federal Minimum Hourly Wage for Farm Workers for the United States [FEDMINFRMWG], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/FEDMINFRMWG, June 3, 2022. Currently, various subgroups such as workers under the age of 20, student workers, workers with disabilities, workers paid at least partially in tips, and "homeworkers making wreaths" have alternative minima or are exempt from the law entirely; see https://www.dol.gov/agencies/whd/minimum-wage/faq accessed July 6, 2022. Adding to the confusion, many states in the U.S. set their own minimum wages, with workers in those states eligible for the maximum of the state and federal minimum wages.

younger workers.[131] They begin by constructing a simple theoretical model of on-the-job training with asymmetric information, heterogeneous workers, and no mechanism for workers to buy training either directly or via lower wages. In their model, no training occurs in the absence of a minimum wage. In contrast, in the presence of a minimum wage, workers with sufficiently low productivity get trained. Such workers have no incentive to move after training because they receive the minimum wage at every firm because it binds for them even with training. This allows the firm such workers initially match with to capture the return on what they pay for training. In their empirical section they find no evidence that minimum wage laws depress training. To our knowledge, they present the first estimates of the effects of the minimum wage on training obtained using actual panel data, as Neumark and Wascher (2001) implement their panel estimator using repeated cross-section data.

Collectively, Grossberg and Sicilian (1999), Neumark and Wascher (2001), and Acemoglu and Pischke (2003) represent a substantial improvement over previous studies because they use measures of actual training rather than inferring it from wage growth. But, of course, *every* study has its limitations. For instance, Neumark and Wascher argue that the Grossberg and Sicilian study suffers from three important limitations. First, they worry that the uncovered sector—the comparison group in their paper—may have less access to training for other reasons, biasing their estimates upward. Second, they note that the starting wage and training are jointly endogenous, making it difficult to interpret estimates of linear models with training as the dependent variable and the starting wage on the right-hand side. Finally, they worry that job complexity, proxied by the number of weeks it takes a new employee to become fully trained and qualified in the position, might soak up much of the variation in training and so leave the analysis without the statistical power to detect effects of modest but meaningful size.

In addition to the reservations expressed in Neumark and Wascher (2001), Gross-

[131] See Sections 3 and 4 for more discussion of the underlying survey questions and the descriptive information on training they provide.

berg and Sicilian add a variable they call related experience, which captures responses to the survey question "How many [years] of experience in jobs that had some application to the position did (name) have before (he/she) started working for your company?" But firms trying to economize on training costs because they must pay their workers the minimum wage might try to hire workers with more related experience. Indeed, the results in Grossberg and Sicilian's Table 3 suggest this might be occurring with men paid at the minimum having (on average) over 300 percent more related experience than workers paid below the minimum wage and about 88 percent more related experience than men paid just above the minimum wage. Things look murkier for women. Women paid at the minimum wage average about 31 percent more related experience than those paid just below the minimum, but women paid just above the minimum average about five percent more related experience than those paid the minimum wage.

As Acemoglu and Pischke (2003, pp. 163) note, the Neumark and Wascher (2001) obtain very large estimates. In their words:

> To see why the effects implied [by] Neumark and Washer's paper are implausibly large, note that their treatment group consists of all young workers. Not all of these workers are affected by the minimum wage, however. Let us assume, quite generously, that all workers earning less than 160% of the minimum are "affected" by the minimum wage. The 160% of the average federal minimum over the period they study is \$5.60, and 40% of workers aged 20-24 are paid below this wage in 1991. Neumark and Washer's estimates imply that formal training among workers aged 20-24 in California (a high minimum wage state) was 3.2 percentage points lower than in states which were subject to the lower federal minimum. This point estimate, then, implies that among affected workers, training will be lower by approximately eight percentage points (i.e. 3.2 percentage points divided by 0.40). The average incidence of training among affected workers in low minimum wage states is 3.0% (much lower than among all workers aged 20-24

> for whom the incidence is 10%). So this estimate implies that introducing California's minimum wage to low minimum wage states should have wiped out all training *two and a half times* among affected workers in these states! Clearly, an implausibly large effect.

A defense of Neumark and Wascher's estimates would note their wide confidence intervals. Their evidence is consistent with smaller but still meaningful negative impacts on training. Alternatively, one can interpret the implausibly large point estimates as a failed specification test that indicates substantive problems with the two comparison groups they employ and/or with the identification assumptions they impose upon them.

Acemoglu and Pischke's study has limitations of its own. By design, NLSY-79 sample members have birthdates in the interval between January 1, 1957, and December 31, 1964. This implies that they range in age from 22 to 28 in calendar year 1986, the reference period for the 1987 survey year in which the NLSY-79 started measuring training. By the end of Acemoglu and Pischke's sample period they range in age from 27 to 33. Some workers do labor at or near the minimum wage at these ages, but such workers constitute a much more strongly selected subset of all workers than do minimum wage workers at younger ages. This selection raises important questions about the external validity of their estimates. More generally, because of this difference in study populations, one should not simply compare the estimates from the two papers. Null effects in Acemoglu and Pischke (2003) can peacefully coexist with negative effects in Newmark and Wascher (2001).

All three papers benefit from using direct measures of training instead of attempting to infer the extent of training from wage profiles. At the same time, our discussion in Section 3, especially our discussion of Barron et al. (1997a), provides strong reasons to think that the available direct measures of training embody substantial measurement error. Hausman's (2001) Iron Law of Measurement Error (i.e., that it causes attenuation bias) then implies bias towards zero for the estimates in all three papers. This renders the null findings in Grossberg and Sicilian (1999) and Acemoglu and Pischke (2003) less compelling (or at least more difficult to interpret) and makes the large effects in Neumark and Wascher (2001) even more puzzling.

Finally, notice the subtle difference between Grossberg and Sicilian (1999) and both Neumark and Wascher (2001) and Acemoglu and Pischke (2003). Grossberg and Sicilian's sample consists of firms offering jobs while the other two papers study samples of the relevant workers. Because the EOPP data samples firms with newly hired workers rather than workers, it over-represents positions with frequent turnover relative to positions with infrequent turnover. If firms economize on training costs by doing less training of workers in high turnover positions, then we would expect lower population training rates for the EOPP sample than for the samples studied in the other two papers.

In our view, none of the issues raised here render any of these three studies, nor the "first-generation" studies that relied on wage growth to proxy for training, completely invalid. Every empirical study has its own distinct set of strengths and weaknesses. We think all these studies deserve positive (though not identical) weights in the reader's casual Bayesian aggregation of the available evidence. With our own evidentiary weights, these earlier studies do not imply a consensus. We now turn to some more recent papers that do not lead us to a consensus either.

Using data from the British Household Panel, Arulampalam et al. (2004b) compare affected workers (measured in two different ways) to unaffected workers before and after the advent of Britain's national minimum wage in a traditional difference-in-differences framework. They find no evidence that the arrival of the national minimum wage reduced training, and some limited evidence that training may have increased for affected workers.

Schumann (2017) studies the effects of the introduction of a sectoral minimum wage in the German construction industry on firms' decisions regarding whether or not to hire any apprentices (the extensive margin) and how many to hire (the intensive margin, both with and without the zeros). His study has the advantage of relying on administrative data, thus avoiding issues of both selective non-response to surveys and the sorts of measurement error lovingly described in Section 3. Using (we sense a theme) a traditional difference-in-differences setup with manufacturing firms as the comparison group, the author finds a decline on the extensive margin in the former East Germany but not in the former West Germany, and declines in the number of apprentices throughout Germany, though with much larger negative effects in the former East Germany. Schumann (2017) attributes the larger effects in the Eastern Länder to the much stronger "bite" of the minimum wage in the east relative to prior wage levels in the construction industry.

In a similar vein, Bellmann et al. (2017) look at short-run response to the creation of the 2015 statutory German minimum wage using the IAB Establishment Panel. They compare establishments at which all employees were above the 2015 minimum wage in the pre-period (the comparison group) to establishments that had some workers below the minimum wage in the pre-period (the treatment group). The data contain measures of whether any training took place at the establishment (their incidence measure) and the fraction of employees trained (their intensity measure). While they find no impact on their incidence measure, they do find robust evidence of a reduction in their intensity measure. In addition, they estimate larger reductions in training in more competitive labor markets, in line with the prediction of Acemoglu and Pischke (2003). Most intriguing of all, they find a stronger decline in training intensity among medium- and high-skill workers than among low-skill workers, suggesting a change in technology at establishments affected by the minimum wage.

All three of these more recent studies provide interesting evidence, and all three should receive positive weight in any cumulative review. They share the advantage of employing more recent data. We can easily imagine that labor markets and skill development technologies have changed sufficiently in recent years in ways that problematize the external validity of the earlier studies reviewed above. At the same time, all three of the recent studies have important limitations that limit their evidentiary weight. Two of the studies use training measures that build on survey reports by firms. While Barron et al. (1997a) find that firms tend to report more training than workers when looking at the same training event, we have reason to think that firms under-report training here. For instance, in the IAB survey that Bellmann et al. (2017) rely on, only 70 percent of establishments report positive training incidence. Taken literally, this implies that in 30 percent of German establishments did not provide *any* training from 2011 to 2015. We find that schwer zu glauben (i.e. hard to believe). The third study, Schumann (2017), provides compelling evidence on German apprenticeship training, but the institutional uniqueness of that system raises concerns about generalizing even to later firm training in Germany.

In our discussion of the literature on the effects of minimum wage laws on firm training, we foregrounded training measurement issues that we expect to cause meaningful attenuation bias in many of the estimates. We have not emphasized two other issues that this narrow literature shares with the much larger literature on the employment effects of minimum wages: policy endogeneity and a general failure to distinguish between short-run and long-run impacts. The studies on training effects all recognize the policy endogenity issue and attempt to deal with it, but one can always quibble with the solutions. We do not have a strong prior about the sign of the bias from policy endogeneity. Most of the papers described in this subsection implicitly offer weighted averages of short-run and longer-run effects, though some, like Schumann (2017), explicitly focus on short-run effects. We usually expect effects to increase with the time allowed for adjustment; for that reason we expect that the available literature on average understates the long-run effects that matter most for policy. In sum, based on our review, we see the literature as indicating modest negative effects of minimum wages on firm training.[132]

6.4 Taxation and training

One important government policy provides an obvious disincentive not only for employer-provided training, but for investments in education as well: the income tax system.

[132]One issue we worry less about in this literature than in others is publication bias. For reasons both theoretical and political, the minimum wage literature allows publications of null findings that would remain forever on the researcher's hard drive in other domains.

Indeed, income taxes provide disincentives to participate in the labor market; among labor market participants, they incentivize workering fewer hours. They also encourage workers to take compensation in (less efficient) non-pecuniary forms. Any progressivity of the tax system makes this "leisure subsidy" higher for the high-skilled.[133]

Even in the "low-tax" environment of the United States, workers face marginal tax rates that provide tremendous disincentives to invest in training. To illustrate, consider a very talented, single worker who earns $100,000 per year. The worker faces a marginal federal income tax rate of 24 percent.[134] But we are not done. Most US states have income taxes (while a few do not). If one takes the five largest states by population, two do not have an income tax (Texas and Florida) and three do (California, New York, and Pennsylvania). If our worker lives in Pennsylvania (which has the median tax rate on her earnings), she faces a state tax rate of 3.07 percent, which raises her full marginal tax rate to 27.07 percent. If she moves to California, her marginal state tax rate rises to 9.3 percent and her combined marginal tax rate increases to 33.3 percent.[135]

Tax rates like these (and the substantially higher rates in many other jurisdictions) surely deter otherwise efficient investments in firm training, and thus likely justify some sort of training subsidy. But the existing literature offers very little in terms of evidence on magnitudes, making this a fruitful area for future research in our view.

[133] Though the implicit tax on means-tested benefits often makes marginal tax rates on very low incomes extremely high as well. For instance, the U.S. is particularly effective at taxing the poor, with marginal tax rates often exceeding 100 percent; see e.g. Kosar and Moffitt (2017).

[134] Our worker must also pay 7.65 percent of her earnings toward Social Security and Medicare benefits for existing retirees, and her employer pays another 7.65 percent. If we assume that the worker implicitly pays the employer share via lower wages, then she faces an overall marginal federal tax rate of 39.3 percent. But our worker will also receive Medicare and Social Security benefits later in her life, and the level of her Social Security benefits will depend directly on her earnings history. Moreover, the value of the Medicare and Social Security benefits ultimately depends on her expected longevity, which makes this problem particularly difficult. We simply assume the expected benefits compensate for the taxes paid, although given her relatively high income, this probably understates her tax bill.

[135] Should she invest in so much training that her annual earnings exceed $164,925 per year, her marginal federal income tax rate would jump to 32 percent, an eight percentage point increase, but the Social Security tax would disappear at the margin.

6.5 Policy toward non-compete clauses

Firms appear to pay a portion of general training costs. How can they expect to recover these costs? Following Becker, one way requires workers to accept lower wages for a period long enough to cover the cost of the training. A disadvantage to this contract is that if the workers find they do not want to work in the relevant job they must nonetheless continue to work until they repay their bond or else directly compensate the firm. Workers quite reasonably worry that firms will use the bond to worsen the non-wage attributes of the position, knowing that workers may find it difficult to quit and repay the training costs. Perhaps not surprisingly, then, the literature contains little empirical evidence of workers paying for a substantial portion of training costs, (e.g., Barron et al. 1989).

On the other side of the labor market, firms know that workers behave strategically in employment relationships too. When firms buy physical capital, they naturally retain the property rights to that capital. In contrast, when they pay for training for a worker, the (human) "capital" remains embedded in the worker. This creates a potential hold-up problem when workers leave the firm: Firms quite reasonably worry that workers may quit after becoming fully trained, and, to add insult to injury, go to work for a competitor. Non-compete agreements ameliorate this incentive problem by having the worker agree not to work for a competitor for some period of time after leaving the training firm. Thus, if workers discover they hate the occupation for which they are training, nothing prevents them from trying a different occupation, but firms do not have to worry about training a workforce for their competitors. To economists, non-compete clauses provide employers with a degree of monopsony power, while also encouraging the provision of productive training. Economic efficiency requires finding the right balance between these competing forces.

To further illuminate the key issues, consider an extreme version of a non-compete that simply binds the worker to the training firm for a period of time long enough for the worker to "work off" the cost of their training. As described by Galenson (1984) and others, such contracts, which create an employment relationship called "indentured servitude", helped to finance the transportation of many credit-constrained workers from Europe to the Americas back in the colonial era. The 13^{th} Amendment to the U.S. Constitution explicitly proscribes involuntary servitude (as well as slavery), which renders contracts that require the worker to remain at the training firm illegal. This generates questions and problems for non-compete contracts like the ones we actually observe in the modern U.S. labor market: How much may non-compete contracts restrict the alternatives of workers without violating the ban on indentured servitude? Can non-competes force workers to relocate in order to stay in the same occupation or industry? How long may a non-compete prevent the worker from working at a competitor? Can they permanently ban workers from an industry? Drawing the line between non-compete clauses and indentured servitude is a crucial issue.

Non-compete agreements conjure up images of employers seeking to protect trade secrets when they hire workers for high-tech jobs. Consistent with such images, Starr et al. (2021) estimate that non-compete agreements cover 35 percent of the workers in computer and mathematical occupations. But they also estimate that 11 percent of building and grounds cleaning and maintenance workers, 11 percent of workers in food preparation and serving related occupations, and 19 percent of workers in protective service occupations labor under non-compete agreements. Indeed, they estimate that 20 percent of all workers with less than a high school degree have a non-compete agreement with their current employer, and that 38 percent of all workers have at some point in their career signed a non-compete agreement. Legally, most U.S. states enforce at least some non-compete agreements, but the exact policies vary greatly across states. For instance, North Dakota and California do not enforce such agreements at all. In contrast, other states enforce such agreements to some degree, not only when workers quit, but when firms dismiss them or lay them off as well. Clearly, non-compete agreements play a much larger role in the U.S. labor market than uninformed intuition might suggest.[136]

[136] Our discussion in this subsection focuses on the U.S. literature. See e.g. Young (2021) for an analysis of a ban on low-income non-competes in Austria as well as pointers to the broader international literature.

Starr (2019) explores the link between training and non-compete agreements using data from the Survey of Income and Program Participation (SIPP). The SIPP includes the following training question:

> "During the past year, has [the respondent] received any kind of training intended to improve skill in one's current or most recent job?"

This question differs from the training questions in other U.S. surveys listed in Table 1.[137] According to Starr (2019), 21 percent of SIPP respondents give a positive response to this question. This compares to the training rate of roughly 43 percent that Neumark and Wascher (2001) calculate based on the CPS training supplement. Differences in the populations of interest in the two studies complicate the comparison as, for example, Neumark and Wascher include 16- and 17-year-olds and omits everyone over the age of 54, both choices that we would expect to drive up the rate. Still, our discussion in Section 3.3, along with the comparison with the CPS, leads us to expect the SIPP to understate the quantity of training. How, and how much, such measurement error might affect the Starr (2019) estimates of the effect of non-compete enforcement on firm training remains an open question.

Starr measures enforcement by constructing an index based on multiple dimensions of non-compete policy at the state level. He then undertakes a difference-in-differences style analysis that implicitly compares occupations with high and low rates of non-competes in states with high and low levels of enforcement. He finds that going from zero to the mean of the enforceabilty index results in a 14 percent increase in training incidence, along with a four percent decline in hourly wages. Increasing enforceability also reduces the "return" to job tenure as well as job turnover. Looking at subgroups, Starr finds a larger effect on hourly wages for workers with less than a graduate degree.

In 2008 Oregon banned the use of non-compete agreements for hourly employees. Lipsitz and Starr (2022) analyze the consequences of that ban using data from CPS Outgoing Rotation Groups (ORG) and both difference-in-differences and synthetic con-

[137] Given the general absence of research on question wording effects for survey measures of training, we wonder what sorts of conversations led to the variation we observe in practice.

trol designs. They find that hourly employees experienced a 2-3 percent increase in earnings as a result of the ban. They scale this up to an increase of about 14 to 21 percent for workers covered by non-compete agreements, while noting that potential importance of spillovers to other workers. The earnings increase appears larger for female workers. They also find that the ban increases job-to-job mobility for hourly workers and increases the proportion of salaried workers, the latter presumably the result of firms seeking to avoid the ban on non-compete agreements.

Theory does not sign the welfare effect of allowing (or not) non-compete agreements. On the one hand, non-competes allow firms to finance a portion of the general training costs, which may benefit credit-constrained, low-skill workers. On the other hand, non-compete agreements appear to increase the monopsony power of firms, which we seldom associate with an increase in economic welfare. Starr et al. (2020) look at the (common) firm practice of having workers sign non-compete agreements even in states that will not enforce them. If all parties were fully informed of their rights, such behavior would be extremely hard to justify. Of course, businesses likely do understand the legalities involved. Indeed, the less generous among us might view such unenforceable agreements as an attempt to take advantage of the less informed workers. The results of Starr et al. feed this admittedly cynical interpretation. They find workers in states that do not enforce non-compete agreements just as likely to cite their signing of a non-compete agreement as a reason for not accepting an offer from a competitor of their current firm.

6.6 Sectoral training programs

After many years of public job training programs usually failing to pass simple cost-benefit tests, the last few years have witnessed a wave of successful training programs that appear to unlock, at least upon occasion, the sorts of returns found in the firm training literature. These programs are called sectoral training programs and they have understandably generated some excitement. Though the financing typically comes from government or non-profit sources, we consider these programs in a chapter on firm training because of their strong focus on employer involvement, particularly in curriculum development.[138]

Katz et al. (2022, KRHS henceforth) provide an expert summary and an analysis of the potential mechanisms underlying the positive impacts. Table 1 of KRHS describes the 10 experimental evaluations of sectoral training programs they consider, including program details and impacts on earnings in the two years after random assignment. In percentage terms, the earnings impacts range from -17.7 percent to 37.7 percent, with a median around 26.7 percent.[139]

KRHS offer two theoretical stories to account for these large effects. The first they term "static (or persistent) inefficiencies in training provision" by which they mean that the monopsony power of firms may lead to the poaching externality we discuss in Sections 5.3, 5.4, and 6.1.3 above. Of course, to the extent the programs provide skills valued by a variety of firms, the standard models considered in Section 5.1 would see workers bearing the costs of this training and reaping all of the gains, though monopsony would mute these gains and thus lead to under-investment in the relevant skills. One could couple a monopsony power story with imperfections in the capital market to explain why less disadvantaged workers acquire the skills these programs provide without the programs, but more disadvantaged workers do not. Such a model would also help account for the relatively large impacts from programs focused on the disadvantaged.

The second story KRHS term "dynamic adjustment and inefficiencies in training provision" by which they essentially mean the ability to pick sectors where (loosely speaking) demand for workers exceeds the current supply. While definitely a possibility, if this mechanism generates the large returns, we see that as bad news for sectoral training programs. Presumably, picking sectors, like picking stocks, involves a great deal of luck. How would the government, or firms for that matter, know that particular providers, such as the four non-profit agencies that run the WorkAdvance programs,

[138]See e.g. Barnow and Spaulding (2019) on this point.

[139]The lone negative two-year earnings impact corresponds to a program where most participants remained in training in year two. Its impacts become positive and substantial in later years.

had the sector-picking skill rather than just having good luck?[140]

In their effort to understand the effectiveness of sectoral training programs, KRHS focus on the WorkAdvance model that aggregates four programs: Per Scholas in New York City (which targeted IT), Towards Employment in Northeast Ohio (which targeted health care and manufacturing), Madison Strategies in Tulsa, Oklahoma (which targeted transportation and manufacturing), and St. Nick's Alliance also in New York City (which targeted environmental remediation).[141] All four WorkAdvance programs impart both general skills and occupation-specific skills, but not, in general, any firm-specific skills. While run by different organizations, KRHS argue for grouping the programs because of their common elements: (1) Specific occupational skills training; (2) Career readiness services; (3) Job development and placement services; (4) Employment retention and career advancement services; and (5) Screening prior to program enrollment. We follow the lead of KRHS and highlight the WorkAdvance programs, but other sectoral training programs share these elements.[142]

Consider each of the program elements in turn. As documented in great detail in e.g. McCall et al. (2016), government programs (and community colleges) have long offered specific occupational training. The meta-analysis of Card et al. (2018) confirms the general sense in the literature that, averaging across both participants and programs, such training has modest but not zero effects on employment and earnings. But these modest impacts look nothing like the very large impacts found in the WorkAdvance studies, so other factors must matter as well.

We emulate KRHS and group the next three program elements together under the heading of "wraparound" services. Some existing evidence sheds light on the potential contributions of these services to the overall impacts. For example, relatively inten-

[140]Indeed, we could follow the ubiquitous practice of interpreting impact estimates not statistically different from zero as equal to zero and conclude that, among the four WorkAdvance programs, only Per Scholas got "lucky" in this sense, as the other three programs generate only statistically insignificant impact estimates. We could do that, but it would be wrong.

[141]Greenberg and Schaberg (2020) offer a deeper dive into the details of the four programs, along with medium-term impact estimates.

[142]For instance, see Roder and Elliot's (2018) report on Quest and Fein and Hamadyk's (2018) report on the Year-Up program.

sive wraparound services provided to community college students via the Accelerated Study in Associates Programs (ASAP) have demonstrated sizeable effects on two-year degree completion, as shown in Gupta (2017) and Miller et al. (2020). Griffith and Rask (2016) document impacts of student services expenditures at four-year colleges. On the other hand, Hamilton and Scrivener (2012) describe the weak findings from the Employment Retention and Advancement Project, which provided related services to welfare recipients and other low-income individuals. Of course, the interventions considered in these other studies do not correspond directly to the services provided to WorkAdvance participants, and considering them in isolation differs from considering them as WorkAdvance combined them. Noting that more research (including analytic literature reviews focused on particular services) would have value here, we think that elements (2), (3) and (4) likely play a small but not zero role in the overall story.

The substantial screening of program applicants prior to enrollment represents perhaps the most distinctive element of the Workadvance programs. Existing public job training programs often do some screening and sorting conditional on enrollment, as when assigning the most job-ready participants to job search assistance rather than to vocational training, but typically do less screening at the enrollment stage. Indeed, Heckman et al. (1996) argue that many public program operators seek out the “hard to serve” among the eligible as most deserving of public resources. In contrast, the WorkAdvance programs engage in strong positive selection on factors, such as prior education and test scores, thought to predict successful engagement with the training. To the extent that existing skills complement rather than substitute for the skills provided in the sectoral training, this screening may account for some or all of the large impacts relative to prior efforts. Moreover, as we documented in Section 4, firms tend to train their strongest workers; in this sense, WorkAdvance mimics private sector practice.

Some casual evidence for a role for screening comes from a comparison of the impacts and control group outcome levels of the Per Scholas and St.Nick’s Alliance WorkAdvance programs, both of which operate in New York City. Per Scholas has both the higher control group earnings levels, indicating stronger screening, and the larger impacts. If screening drives the sectoral training impacts, this has two important implications: First, the program may still be worth doing, but an important aspect of its value may come from certifying non-traditional training applicants as likely to succeed. Second, it suggests that scaling-up the programs by serving less heavily selected applicants will reduce their earnings effects.

Several other factors muddy the explanatory waters even further. First, some recent sectoral programs fail to provide strong positive impacts on labor market outcomes. A leading example is the Health Professional Opportunity Grants (HPOG) program evaluated in Peck et al. (2022). Second, one might worry about spillovers, either via displacement or via skill prices, that lead the partial equilibrium experimental impact estimates to overstate the overall economic impacts. KRHS offer some reasons for expecting modest spillover effects in the sectoral training context. Ultimately, though, their extent represents an empirical question. Unfortunately, our best evidence on spillovers comes from programs aimed primarily at improving job search, e.g. Lise et al. (2004), Crépon et al. (2013), and Gautier et al. (2018), or from the context of four-year degrees as in Heckman et al. (1998). The literature awaits a more directly relevant empirical study. Finally, perhaps these programs arise because of genuine shortages of particular sets of skills in specific local labor markets. Their initial cohorts realize large impacts as they ameliorate the shortage, but those impacts may not recur in later cohorts precisely because the initial cohorts have "solved" the problem.

Overall, while we share some of the enthusiasm for sectoral programs expressed in KRHS and in the policy literature, we think that a substantial amount of research remains to understand the mechanisms that underlie their performance and so constitute the "secret sauce" that differentiates them from earlier, less successful programmatic efforts. Sorting empirically among the potential mechanisms is critical for thinking about the nature and size of the benefits from scaling up these sectoral programs. We expect this research task to keep the "Workforce Industrial Complex" busy for quite some time.

7 Conclusion

Our study of firm training, which included reviews of several components of the literature as well as some fresh empirical work, yields broad conclusions related to measurement, importance, heterogeneity, impacts, theory, and policy. We summarize our conclusions under each heading in turn.

Brown (1989a), in his long-ago survey of the training literature, wrote that "The current method of collecting information about training does not give any reason to hope that we will know much more about the subject in ten years than we do today." Though we have some sympathy with his forecast, we think we have learned something about firm training over the past three decades, and even more about the challenges associated with measuring firm training. Our review of the literature makes clear that seemingly minor differences in survey questions, or differences in the routing that leads to the questions, or whether the worker or a representative of the firm answers the question, can lead to quite different estimates of the incidence, duration and determinants of training. For example, one in three Dutch workers (in the 1990s) participated in training according to IALS but only one in twenty did so according to the contemporary ECHP. Further research that attempts to explain the variation in incidence measures across surveys (and subgroups, countries, and time) would add great value to the literature. So would additional research based on administrative training data and research quantifying the nature and importance of casual training on the internet via videos (e.g. via YouTube™videos).

Under the second heading, the literature shows that the empirical importance of firm training, conceived of as its contribution to the overall stock of human capital, clearly exceeds that of formal schooling. Society quietly spends more time and other resources on firm training than on formal schooling (despite vast expenditures on the latter). More than one in ten American private-sector workers report participating in formal training each year. Two in three report participating in informal training. The OECD (at least the subset of it represented in the PIAAC) boasts only modestly lower participation rates: 0.13 vs 0.09 for formal training and 0.66 vs 0.54 for informal training. Employers finance much of this training. Indeed, our review suggests that in little over a decade an average worker receives human capital investments by their employer equivalent to four years' worth of college going. Yet, for a variety of reasons, data availability issues and naval-gazing by academics among them, the size of the literature on formal schooling exceeds that on firm training by several orders of magnitude. We think this ratio signals a costly misallocation of research activity at the margin.

Under the third heading, we document that the nature and intensity of firm training vary tremendously on multiple dimensions. We began the chapter by discussing four broad categories of workplace training in Section 2: formal training, informal training, casual training and learning by doing. The studies reviewed in Section 4 showed that conditional on category, training spells vary tremendously in the duration, with the standard deviation typically far exceeding the mean. In addition to variation in duration, training also varies widely in content, from short courses designed to introduce new software, to soft skills courses that aim to improve the workplace environment, to annual video reviews of privacy restrictions on sensitive data. This rampant heterogeneity helps generate the measurement difficulties recounted under the first heading and also some of the difficulties in obtaining meaningful estimates of the impact of firm training considered under the next heading.

Under the fourth heading, obtaining plausible estimates of the causal effect of firm training remains a challenge for several reasons. First, and most obviously, existing surveys likely do not collect data on all of the worker characteristics that affect productivity (and thus wages) and training participation. This means trouble, especially for studies that rely on conditional independence assumptions to justify their causal claims. Confounding treatments also matter in this context. For instance, the literature clearly reveals a positive correlation between receipt of formal and informal training. Estimating the impact of one without taking account of the other will lead to over-estimation of its causal effect. In the same vein, it is remarkable how little we still know about serial correlation in training incidence, mostly because we tend to measure training via cross-sectional surveys. If training incidence positively correlates over time, as we suspect it does, then many cross-sectional impact estimates likely confound the cumulative influence of past training with the causal effect of current period training. Finally, the training context offers opportunities for reverse causation: for example, if senior management sees managerial potential in Susan and equips her with two weeks of managerial training before promoting her, claiming the two weeks of training as the sole driver of the wage increase associated with her promotion does a disservice to the true causal mechanism. In short, the literature on firm training has plenty of room for better data on workers characteristics as well as on training, for improved institutional knowledge, and for more compelling sources of variation in training receipt.

Under the fifth heading, we bemoan the general dearth of new theory around firm training, with the important exceptions of Lazear (2009) and Cavounidis and Lang (2020). The widely-cited papers of Acemoglu and Pischke (1998, 1999a, 1999b, 2003) have not attracted as much response and exploration as their citation counts suggest they should, particularly given their mixed empirical performance, as illustrated in e.g. Section 5.4.3. They provided a needed response to the failure of certain aspects of the models in the literature initiated by Becker (1964) and Mincer (1962), such as the failure of the empirical literature to find differences in starting wages associated with firm training. But many unsolved or poorly solved puzzles related to firm training remain. One very important example: How have changes in job search technology changed the relationships between wages, training, and experience?

We also think the literature has paid too little attention to some older theoretical developments. Consider, for example, the links between search, training, and relevant experience highlighted in Sections 5.1 and 5.6. Given positions that require a set of skills not taught at schools or universities, firms face two choices. They may provide existing workers with the training and experience necessary to master those skills. Alternatively, they may hire workers who already possess the necessary training and experience for success in those positions. The second alternative requires that firms search more (or at least differently in a directed search sense) and that they offer higher starting wages. This simple story highlights the important link between search costs and training costs at the firm level. A series of papers by John Barron and co-authors in the 1980s draws out the importance of this relationship, yet the current literature largely ignores it. Instead, it should update this line of work to reflect the fundamental changes in search technology in the intervening decades and to incorporate the firm's choice of work technology.

Finally, under the sixth heading, the literature does not offer up much in the way of evidence-based policy prescriptions, other than the straightforward case for modest tax subsidies to undo the effects of proportional taxes on labor earnings. It does provide limited evidence of interactions between training and policies in other domains, such as limitations on non-compete agreements and minimum wages, with the directions matching what simple theory would predict. It also provides cautionary tales regarding attempts to promote firm training that mostly lead to reclassification of less formal training into more formal training or massive "deadweight" as taxpayers pay for infra-marginal training. More broadly, we find no evidence of any income gradient in the take-up of formal training in either the US or the broader OECD using the newest data available, suggesting a limited role for liquidity constraints. That informal training increases with initial schooling and with income suggests to us the importance of selection on "ability" as a driver of participation in firm training.

In closing, we wrote this chapter in an attempt to summarize and reinvigorate the relatively moribund literature on firm training. We look forward to a more rapid pace of empirical and theoretical development around firm training in the years to come.

8 References

Abraham, K.G., Farber, H.S., 1987. Job duration, seniority, and earnings. American Economic Review 77, 278-97.

Abramovsky, L., Battistin, E., Fitzsimons, F., Goodman, A., Simpson, H., 2011. Providing employers with incentives to train low-skilled workers: Evidence from the UK Employer Training Pilots. Journal of Labor Economics 29, 153-193.

Acemoglu, D., 1997. Training and innovation in an imperfect labour market. Review of Economics Studies 64, 445-64.

Acemoglu, D., Pischke, J.S., 1998. Why do firms train? Theory and evidence. Quarterly Journal of Economics 113, 79-119.

Acemoglu, D., Pischke, J.S., 1999a. The structure of wages and investment in general training. Journal of Political Economy 107, 539-572

Acemoglu, D., Pischke, J.S., 1999b. Beyond Becker: Training in imperfect labour markets. Economic Journal 109, F112-F142.

Acemoglu, D., Pischke, J.S., 2003. Minimum wages and on-the-job training. Research in Labor Economics 22, 159-202.

Aigner, D.J., 1973. Regression with a binary independent variable subject to errors of observation. Journal of Econometrics 1, 49-60.

Adua, L. and J.S. Sharp, 2010. Examining survey participation and response quality: The significance of topic salience and incentives. Survey Methodology, 36, 95-109.

Akerlof, G. A., 1970. The market for 'lemons': Quality uncertainty and the market mechanism. Quarterly Journal of Economics 84, 488-500.

Albrecht, J., Jovanovic, B., 1986. The efficiency of search under competition and monopsony. Journal of Political Economy 94, 1246-1257.

Almeida, R., Carneiro, P., 2009. The return to firm investments in human capital. Labour Economics 16, 97-106.

Altonji, J.G., Shakotko, R.A., 1987. Do wages rise with job seniority? Review of Economic Studies 54, 437-59.

Altonji, J.G., Williams, N., 1997. Do wages rise with seniority? A reassessment. NBER Working paper No. 6010.

Altonji, J.G., Williams, N., 2005. Do wages rise with job seniority? A reassessment. Industrial and Labor Relations Review 58, 370-97.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., Schady, N., 2016. Teacher Quality and Learning Outcomes in Kindergarten. Quarterly Journal of Economics 131, 1415–1453.

Arulampalam, W., Booth, A.L., 2001. Learning and earning: Do multiple training events pay? A decade of evidence from a cohort of young British men. Economica 68, 379-400.

Arulampalam, W., Booth, A. L., Bryan, M. L., 2004a. Training in Europe. Journal of the European Economic Association 2, 346-360.

Arulampalam, W., Booth, A. L., Bryan, M. L., 2004b. Training and the new minimum wage. Economic Journal 114, c87-c94.

Arulampalam, W., Booth, A. L., Bryan, M. L., 2010. Are there asymmetries in the effects of training on the conditional male wage distribution? Journal of Population Economics 23, 251-272.

Barnow B.S., Smith, J.A., 2016. Employment and training programs, in: Moffit, R.A. (Ed.), Economics of means-tested transfer programs in the United States Vol III, University of Chicago Press, Chicago, pp. 127-234

Barnow, B.S., Spaulding, S., 2019. Contracting for employment and training programs, in: Shick, R.A., Martin, L.L. (Eds.), Human Services Contracting: A Public Solutions Handbook. Routledge, New York, pp. 78-94.

Barron, J.M., Berger, M.C., Black, D.A., 1997a. How well do we measure training? Journal of Labor Economics 15, 507-528.

Barron, J.M., Berger, M.C., Black, D.A., 1997b. On-the-Job Training. W.E. Upjohn Institute for Employment Research, Kalamazoo.

Barron, J.M., Berger, M.C., Black, D.A., 1999. Do workers pay for on-the-job training? Journal of Human Resources 34, 235-252.

Barron, J.M., Bishop, J., 1985. Extensive search, intensive search, and hiring costs: New evidence on employer hiring activity. Economic Inquiry 23, 363-82.

Barron, J.M., Bishop, J., Dunkelberg, W.C., 1985. The interviewing and hiring of new employees. Review of Economics and Statistics 67, 43-52.

Barron, J.M., Black, D.A., Loewenstein, M.A., 1987a. Employer size: The implications for search, training, capital investment, starting wages, and wage growth. Journal of Labor Economics 5, 76-89.

Barron, J.M., Black, D.A., Loewenstein, M.A., 1989. Job matching and on-the-job training. Journal of Labor Economics 7, 1-19.

Barron, J.M., Black, D.A., Loewenstein, M.A., 1993. Gender Differences in Training, Capital, and Wages. Journal of Human Resources 28, 343-364.

Barron, J.M., Loewenstein, M.A., 1985. On employer-specific information and internal labor markets. Southern Economic Journal 52, 431-445.

Barron, J.M., Loewenstein, M.A., Feuss, S.M., 1987b. Further Analysis of the Effect of Unions on Training. Journal of Political Economy 95, 632-640.

Bassanini, A., Booth, A., Brunello, G., de Paola, M., Leuven, E., 2007. Workplace Training in Europe, in: Brunello, G., Garibaldi, P., Wasmer E. (Eds.), Education and Training in Europe. Oxford University Press, Oxford, pp. 146-342.

Bassanini, A., Brunello, G., 2011. Barriers to entry, deregulation and workplace training: A theoretical model with evidence from Europe. European Economic Review 55, 1152-76.

Battistin, E., De Nadai, M., Sianesi, B., 2011. Misreported schooling, multiple measures and returns to educational qualifications. Journal of Econometrics 181, 136-150.

Baum, S., Holzer, H., Luetmer, G., 2021. Should the Federal Government Fund Short-

Term Postsecondary Certificate Programs? IZA Discussion Paper No 14109.

Becker, G., 1962. Investment in Human Capital: A Theoretical Analysis. Journal of Political Economy 70, 9-49.

Becker, G., 1964. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. The University of Chicago Press, Chicago.

Becker, G. 1965. A Theory of the Allocation of Time. Economic Journal 75, 493–517.

Becketti, S., Gould, W., Lillard, L., Welch, F. 1998. "The Panel Study of Income Dynamics after fourteen years: An evaluation. The Journal of Human Resources 6, 472-492.

Bellmann, L., Bossler, M., Gerner, H.D., Hübler, O., 2017. Training and minimum wages: first evidence from the introduction of the minimum wage in Germany. IZA Journal of Labor Economic 6, 1-22.

Belly, P., Lochner, L., 2007. The changing role of family income and ability in determining educational achievement. Journal of Human Capital 1, 37-89.

Belson, W. A., 1986. Validity in survey research with special reference to the techniques of intensive interviewing and progressive modification for testing and constructing difficult or sensitive measures for use in Survey Research: A Report. Gower Publishing Company, Limited.

Bender, S., Fertig, M., Görlitz, K., Huber, M., Schmucker, A. 2009. WeLL—Unique linked employer-employee data on further training in Germany. Journal of Applied Social Science Studies 129, 637-643.

Ben-Porath, Y., 1967. The production of human capital and the life cycle of earnings. Journal of Political Economy 75, 352-65

Benson, A., 2013. Firm-sponsored general education and mobility frictions: Evidence from hospital sponsorship of nursing schools and faculty. Journal of Health Economics 32, 149-59.

Berger, D., Herkenhoff, K., Mongey, S., 2022. Labor market power. American Economics Review 112, 1147-93.

Betcherman, G., Leckie, N., McMullen, K., 1997. Developing skills in the Canadian Workplace. The results of the Ekos Workplace Training Survey. Canadian Policy Research Network Study No. W02.

Black, D.A., Loewenstein, M.A., 1997. Dismissals and match-specific rents. Labour Economics 4, 325-340.

Black, D.A., Noel, B.J., Wang, Z., 1999. On-the-job training, establishment size, and firm size: Evidence for economies of scale in the production of human capital. Southern Economic Journal 66, 82-100.

Black, D.A, Sanders, S.G., Taylor, L., 2003. Measurement of higher education in the Census and Current Population Survey. Journal of the American Statistical Association 98, 545-554.

Black, S., Lynch, L., 1996. Human-capital investments and Productivity. American Economic Review 86, 263-267.

Blinder, A.S., Krueger, A.B., 1996. Labor turnover in the USA and Japan: A tale of two countries. Pacific Economic Review 1, 27-57.

Bloom, H., Orr, L., Bell, S. Cave, G., Doolittle, F., Lin, W., Bos, J., 1997. The benefits and costs of JTPA Title II-A programs: Key findings from the National Job Training Partnership Act study. Journal of Human Resources 52, 549-576.

Blundell, R., Dearden, L., Meghir, C., 1996. The determinants and effects of work-related training in Britain. The Institute for Fiscal Studies.

Böhning, D., Mahlzahn, U., Dietz, E., Schlattmann, P., 2002. Some general points in estimating heterogeneity variance with the DerSimonian-Laird estimator. Biostatistics 3, 445-457.

Borja, G.J., 1999. The Economic Analysis of Immigration, in: Ashenfelter, O.C., Card, D. (Eds.), Handbook of Labor Economics vol. 3, Elsevier Science B.V., pp. 1697-1760.

Bound, J., Brown, C., Mathiowetz, N., 2001. Measurement error in survey data, in: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics vol. 5, Elsevier Science B.V., pp. 3705-3843.

Booth, A.L., Bryan, M.L., 2005. Testing some predictions of human capital theory: New training evidence from Britain. Review of Economics and Statistics 87, 391-394.

Booth, A.L., Bryan, M.L., 2007. Who pays for general training in private sector Britain? Research in Labor Economics 26, 85-123.

Brown, C., 1989a. Empirical evidence on private sector training. U.S. Department of Labor Background Paper 7a.

Brown, J.N., 1989b. Why do wages increase with tenure? American Economic Review 79, 971-991.

Brunello, G., Comi, S. L., Sonedda, D., 2012. Training subsidies and the wage returns to continuing vocational training. Evidence from Italian regions. Labour Economics 19, 361-372.

Brunello, G., Wruuck, P., 2020. Employer provided training in Europe: Determinants and obstacles. EIB Working Papers No. 2020/03.

Buck, N., Burton, J., Laurie, H., Lynn, P., Uhrig, S., 2006. Quality profile: British household panel survey – version 2.0: Waves 1 to 13: 1991-2003. ISER report, Essex.

Bundesinstitut für Berufsbildung in cooperation with Statistics Finland, FÁS Training and Employment Authority, 3s Research Laboratory, and Statistics Sweden, 2005. Work-package 5: Survey Guidelines. Paper 2: Mode of data collection. Development of a methodology for a long-term strategy on the Continuing Vocational Training Survey (CVTS). Eurostat.

Bushway, S., Johnson, B., Slocum, L., 2007. Is the Magic Still There? The Use of the Heckman Two-Step Correction for Selection Bias in Criminology. Journal of Quantitative Criminology 23, 151–178.

Calónico, S., Smith, J.A., 2017. The women of the National Supported Work Demonstration. Journal of Labor Economics 35, S65-S97.

Cameron, S., Taber, C. 2004. Estimation of educational borrowing constraints using returns to schooling. Journal of Political Economy 112, 132-182.

Card, D., 2022. Who set *your* wage? American Economic Review 112, 1075-90.

Card, D., Kluve, J., Weber, A., 2018. What Works? A Meta-analysis of Recent Labor Market Program Evaluations. Journal of the European Economic Association 16, 894-931.

Carneiro, P., Heckman, J.J., 2002. The evidence on credit constraints in post-secondary schooling. Economic Journal 112, 705-734.

Carmichael, L., 1983a. Firm-specific capital and promotion ladders. Bell Journal of Economics 14, 251-258.

Carmichael, L., 1983b. Does rising productivity Explain seniority rules for layoffs? American Economic Review 73, 1127-1131.

Carmichael, L., 1985. Wage profiles, layoffs, and specific training: Comment. International Economic Review 26, 747-751.

Carrington, W.J., 1993. Wage losses for displaced workers: Is it really the firm that matters? Journal of Human Resources 28, 435-462.

Carruthers, C.K. Jepsen, C., 2021. Vocational Education: An international perspective, in McCall, B.P (Ed.), The Routledge Handbook of the Economics of Education, Routledge: London, 343-380.

Cavounidis, C., Lang, K., 2020. Ben-Porath meets Lazear: Microfoundations for dynamic skill formation. Journal of Political Economy 128, 1405-1435.

Chang, C., Wang, Y., 1996. Human capital investment under asymmetric information: The Pigovian conjecture revisited. Journal of Labor Economics 14, 505-519.

Chiswick, B., Miller, P. 2015. Handbook of the Economics of International Migration Vol 1, Elsevier Science B.V.

Clemens, J. 2021. How do firms respond to minimum wage increases? Understanding the relevance of non-employment margin. Journal of Economic Perspectives 35, 51–72.

Cortés, P., Pan, J., forthcoming. Children and the Remaining Gender Gaps in the Labor Market. Journal of Economic Literature.

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., Zamora, P., 2013. Do labor market policies have displacement effects? Evidence from a clustered randomized design. Quarterly Journal of Economics 128, 531-580.

Crépon B., Ferracci M., Jolivet G. and van den Berg G. J. 2009. Active labour market policy effects in a dynamic setting. Journal of the European Economic Association 7(2–3), 595–605.

Croce, G., di Porto, E., Ghignoni, E., Ricci, A., 2017. Agglomeration and workplace training: knowledge spillovers versus poaching. Regional Studies 11, 1635-1651.

Dauth, C., 2020. Regional discontinuities and the effectiveness of further training subsidies for low-skilled employees. Industrial Labor Relations Review 73, 1147-1184.

Dauth, C., Toomet, O., 2016. On government-subsidized training programs for older workers. Labour 30, 371-392.

Davidoff, T., 2006. Labor income, housing prices, and home-ownership. Journal of Urban Economics 52, 209-235.

De Grip, A., Sauermann, J., 2012. The effects of training on own and co-worker productivity: Evidence from a field experiment. Economic Journal 122, 376-399.

Dearden, L., Reed, H., Van Reenen, J., 2006. The impact of training on productivity and wages: Evidence from British panel data. Oxford Bulletin of Economics and Statistics 68, 397-421.

Dostie, B., 2015. Do train-or-pay schemes really increase training levels? Industrial Relations 54, 240-255.

England, P., 1982. The failure of human capital theory to explain occupational sex segregation. Journal of Human Resources 17, 358-370.

England, P., 1985. Occupational segregation: rejoiner to Polachek. Journal of Human Resources 20, 441-443.

Faberman, R.J., Mueller, A.I., Şahin, A., Topa, G., 2022. Job search among employed and non-employed. *Econometrica* 90(4) 1743-79.

Fein, D., Hamadyk, J., 2018. Bridging the Opportunity Divide for Low-Income Youth: Implementation and Early Impacts of the Year Up Program, OPRE Report Number 2018-65, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Fisher, I., 1930. Theory of Interest. The Macmillan Company, New York.

Fitzgerald, J., Gottschalk, P., Moffitt, R. 1998. An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics. Journal of Human Resources 33, 251-299.

Fleisher, B.M., 1981. Minimum Wage Regulation in Retail Trade. American Enterprise Institute, Washington DC.

Fortson, K., Rotz, D., Burkander, P., Mastri, A., Schochet, P., Rosenberg, L., McConnell, S., 2017. Providing Public Workforce Services to Job Seekers: 30-month Impact Findings on the WIA Adult and Dislocated Worker Programs. Mathematica Policy Research, Washington, DC.

Frazis, H., Gittlemann, M., Joyce, M., 2000. Correlates of training: An analysis using both employer and employee characteristics. Industrial and Labor Relations Review 53, 443-462.

Frazis, H., Loewenstein, M.A., 2003. Estimating linear regressions with mismeasured, possibly endogenous, binary explanatory variables. Journal of Econometrics 117, 151-178.

Frazis, H., Loewenstein, M.A., 2005. Reexamining the returns to training: Functional form, magnitude, and interpretation. Journal of Human Resources 40, 453-476.

Friedman, M., 1962. Capitalism and Freedom. University of Chicago Press, Chicago.

Frölich, M., Lechner, M. 2010. Exploiting regional treatment intensity for the evaluation of labor market policies. Journal of the American Statistical Association 105(491), 1014–29.

Galenson, D.W., 1984. The rise and fall of indentured servitude in the Americans: An economic analysis. Journal of Economic History 44, 1-26.

Gardner, J.M., 1995. Worker displacement: A decade of change. Monthly Labor Review April, 45-57.

Gathmann, C., Schönberg, U. 2010. How general is human capital? A task-based approach. Journal of Labor Economics 28, 1-49

Gautier, P., Muller, P., van der Klaauw, B., Rosholm, M., Svarer, M., 2018. Estimating equilibrium effects of job search assistance. Journal of Labor Economics 36, 1073-1125.

Goglio, V., Meroni, E.C., 2014. Adult Participation in Lifelong Learning. Technical briefing. Joint Research Centre of the European Commission publication No. JRC92330.

Görlitz, K., 2010. The effect of subsidizing continuous training investments – Evidence from German establishment data. Labour Economics 17, 789-798.

Görlitz, K., 2011. Continuous training and wages: An empirical analysis using a comparison-group approach. Economics of Education Review 30, 691-701.

Görlitz, K., Tamm, M., 2017. Information, financial aid and training participation: Evidence from a randomized field experiment. Labour Economics 47, 138-148.

Grossman, M., 2000. The Human Capital Model, in: Culyer, A.J., Newhouse, J.P. (Eds.), Handbook of Health Economics vol. 1, Elsevier Science B.V., pp. 347-408.

Greenberg, D.H., Robins, P.K., 2008. Incorporating nonmarket time into benefit-cost analyses of social programs: An application to the self-sufficiency project. Journal of Public Economics 92, 766-794.

Greenberg, D.H., Schaberg, K., 2020. Long-term effects of a sectorial advancement strategy: Costs, benefits, and impacts from the WorkAdvance demonstration. Manpower Demonstration Research Corporation, New York.

Griffith, A., Rask, K., 2016. The effect of institutional expenditures on employment outcomes and earnings. Economic Inquiry 54, 1931–1945.

Grossberg, A.J., Sicilian, P., 1999. Minimum wages, on-the-job training, and wage growth. Southern Economic Journal 65, 539-56.

Guarino, C., Maxfield, M., Reckase, M., Thompson, P., Wooldridge, J., 2015. An evaluation of empirical Bayes's estimation of value-added teacher performance measures. Journal of Educational and Behavioral Statistics 40, 190–222.

Gupta, H., 2017. The Power of Fully Supporting Community College Students: The Effects of CUNY's Accelerated Study in Associates Programs after Six Years. Manpower Demonstration Research Corporation, New York.

Hamilton, G., Scrivener, S., 2012. Increasing Employment Stability and Earnings for Low-Wage Workers: Lessons from the Employment Retention and Advancement (ERA) Project. OPRE Report 2012-19, Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.

Hanson, M., 2022. Average Cost of College and Tuition. EducationData.org, January 27, 2022.

Hashimoto, M., 1981. Firm-Specific Human Capital as a Shared Investment. American Economic Review 71, 475-81.

Hashimoto, M., 1982. Minimum wage effects on training on the job. American Economic Review 72, 1070-1087.

Hashimoto, M., Yu., B.T., 1980. Specific Capital, Employment Contracts, and Wage Rigidity. Bell Journal of Economics 11, 536-49.

Hausman, J., 2001. Mismeasured variables in econometric analysis: problems from the right and problems from the left. Journal of Economic Perspectives 15, 57-67.

Heberling, A., Baumgartner, 1978. Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. American Sociological Review 43, 447-462.

Heckman, J.J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic Social Measurement 5, 475-492.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47, 53-161.

Heckman, J.J., 1981. Heterogeneity and state dependence, in: Rosen, S. (Ed.) Studies in the Labor Market, University of Chicago Press, Chicago, pp. 91-140.

Heckman, J.J, Hohmann, N., Smith, J.A., Khoo, M., 2000. Substitution and dropout bias in social experiments: A study of an influential social experiment. Quarterly Journal of Economics 115, 651-694.

Heckman, J.J. Hotz, V.J., 1989. Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training. Journal of the American Statistical Association. 84(408), 862-874.

Heckman, J.J., Lochner, L.J., Cossa, R., 2002. Learning-By-Doing vs. On-the-Job Training: Using Variation Induced by the EITC to Distinguish Between Models of Skill Formation. NBER Working Paper No. 9083.

Heckman, J.J., Lochner, L.J., Smith, J.A., Taber, C. 1997. The effects of government policy on human capital investment and wage inequality. Chicago Policy Review 1, 1-40.

Heckman, J.J, Lochner, L.J., Taber, C., 1998. General equilibrium treatment effects: A study of tuition policy. American Economic Review 88, 381-386.

Heckman, J.J., Lochner, L.J., Todd, P.E., 2003. Fifty years of Mincer earnings regressions. NBER Working Paper No. 9732.

Heckman, J.J., Smith, J.A., Taber, C., 1996. What do bureaucrats do? The effects of performance standards and bureaucratic preferences on acceptance into the JTPA program, in: Libecap, G. (Ed.), Advances in the Study of Entrepreneurship, Innovation and Economic Growth: Reinventing Government and the Problem of Bureaucracy vol 7, JAI Press, Greenwich, pp. 191-217.

Heiler, P., Knaus, M. 2022. Effect or treatment heterogeneity? Policy evaluation with aggregated and disaggregated treatments. arXiv:2110.01427v2 [econ.EM].

Hidalgo, D., Oosterbeek, H., Webbink, D., 2014. The impact of training vouchers on low-skilled Workers. Labour Economics 31, 117-128.

Hobijn, B., Şahin, A., 2009. Job-finding and separation rates in the OECD. Economic Letters 104, 107-111.

Holzer, H.J., 1996. What employers want: Job prospects for less-educated workers. Russell Sage Foundation, New York.

Holzer, H.J., Block, R.N., Cheatham, M., Knott, J.H., 1993. Are training subsidies for firms effective? The Michigan experience. Industrial and Labor Relations Review 46, 625-636.

Holzer, H.J., Reaser, J., 1999. Firm-level training for newly hired workers: Its determinants and effects. Research in Labor Economics 18, 377-402.

Hotz, V.J., Scholz, K., 2002. Measuring employment and income outcomes for low income populations with administrative and survey data, in Ver Ploeg, M., Moffitt, R., Citro, C., (Eds.), Studies of Welfare Populations: Data Collection and Research Issues, National Academy Press, Washington, DC., pp. 275-315.

Hui, T.S., Smith, J.A., 2002. The determinants of participation in adult education and training in Canada. Munich Personal RePEc Archive Paper No. 17,998.

Jacobson, L.S., LaLonde, R.J., Sullivan, D.G., 1993. Earnings losses of displaced workers. American Economics Review 83, 685-709.

Jäger, S., Heining, J., 2019. How substitutable are workers? Evidence from worker deaths. Munich PErsonal RePEc Archive Paper No. 109,757.

Jahromi, A., Callaway, B. 2022. Heterogeneous effects of job displacement on earnings. Empirical Economics 62, 213–245.

Jia, N., Molloy, R., Smith, C., Wozniak, A., forthcoming. The Economics of Internal Migration: Advances and Policy Questions. Journal of Economic Literature.

Johnson, W.R., 1978. A theory of job shopping. Quarterly Journal of Economics 92, 261-78.

Jovanovic, B., 1979a. Job matching and the theory of turnover. Journal of Political Economy 87, 972-990.

Jones, D. C., Kalmi, P., Kauhanen, A., 2012. The effects of general and firm-specific training on wages and performance: Evidence from banking. Oxford Economic Papers 64, 151-175.

Kambourov, G., Manovskii, I. 2009. Occupational specificity of human capital. International Economic Review 50, 63–115.

Katz, E., Ziderman, A., 1990. Investment in general training: the role of information and labour mobility. Economic Journal 100, 1147-1158.

Katz, L., Roth, J., Hendra, R., Schaberg, K., 2022. Why do sectoral employment programs work? Lessons from WorkAdvance. Journal of Labor Economics 40, S249-S291.

Kemple, J., Doolittle, F., Wallace, J., 1993. The national JTPA study: Final implementation report. Manpower Demonstration Research Corporation, New York.

Kletzer, L.G., 1989. Returns to seniority after permanent job loss. American Economic Review 79, 536-43.

Kline, P.M., Rose, E.K., Walters, C.R., 2022. Systemic discrimination among large US employers. Quarterly Journal of Economics 137, 1963-2036.

Konings, J., Vanormelingen, S., 2015. The impact of training on productivity and Wages: Firm-Level Evidence. Review of Economics and Statistics 97, 485-497.

Kosar, G., Moffitt, R.A., 2017. Trends in cumulative marginal tax rates facing low-income families, 1997-2007. Tax Policy and the Economy 31, 43-70.

Kristensen, N., Skipper, L., Smith, J.A, 2022. How well do we measure work-related training? mimeo.

Krolikowski, P. 2018. Choosing a control group for displaced workers. Industrial and Labor Relations Review 71, 1232–1254.

Krueger, A., Rouse, C., 1998. The effect of workplace education on earnings, turnover, and job performance. Journal of Labor Economics 16, 61-94.

Kuhn, P., Mansour, H., 2014. Is internet job search ineffective? Economic Journal 124, 1213-33.

Kuhn, P., Yu., L, 2021. How costly is turnover? Evidence from retail. Journal of Labor Economics 39, 461-95

Kuziemko, I., Pan,J., Shen, J., Washington, E. 2018. The mommy effect: do women anticipate the employment effects of motherhood. NBER Working Paper No. 24740.

LaLonde, R. 1986. Evaluating the econometric evaluations of training programs using experimental data. American Economic Review 76, 604–20.

Lamadon, T., Mogstad, M., Setzler, B., 2022. Imperfect competition, compensating differentials, and rent sharing in the US labor Market. American Economic Review 112, 169-212.

Lazear, E.P., 2009. Firm-specific human capital: A skill-weights approach. Journal of Political Economy 117, 914-940.

Lazear, E.P., Miller, F.H., 1981. Minimum wage versus minimum compensation, In: Report of the Minimum Wage Study Commission. U.S. Government Printing Office Washington DC.

Lerman, R. I., McKernan, S.-M., Riegg, S., 1999. Employer-Provided Training and Public Policy. Available at the U.S. Dept of Education, Educational Resources Information Center (ERIC).

Leuven, E., Oosterbeek, H., 1999. The demand and supply of work-related training: evidence from four countries. Research in Labor Economics 18, 303-330.

Leuven, E., Oosterbeek, H., 2004. Evaluation the effect of tax deductions on training, Journal of Labor Economics 22, 461-488.

Leuven, E., Oosterbeek, H., 2008. An alternative approach to estimate the wage returns to private-sector training. Journal of Applied Econometrics 23, 423-434.

Lewbel, A. 2019. The identification zoo. Journal of Economic Literature 57, 835-903.

Lillard, L., Tan, H.W., 1986. Private sector training: Who gets it and what are its effects? Rand Monograph No. R-3331-DOL/RC, Santa Monica.

Lipsitz, M., Starr, E., 2022. Low-wage workers and the enforceability of noncompete agreements. Management Science 68, 143-70.

Lise, J., Seitz, S., Smith, J.A., 2004. Equilibrium policy experiments and the evaluation of social Programs. NBER Working Paper No. 10283.

Lochner, L., Monge-Naranjo, A., 2012. Credit constraints in education. Annual Review of Education 4, 225-56.

Loewenstein, M.A., Spletzer, J.M., 1999a. Formal and informal training: Evidence from the NLSY. Research in Labor Economics 18, 403-438.

Loewenstein, M.A., Spletzer J.M., 1999b. General and specific training. Evidence and implications. Journal of Human Resources 34, 710-733.

Lynch, L.M. 1992. Private-sector training and the earnings of young workers. American Economic Review 82, 299-312.

Lynch, L.M. 2000. Reorienting training policies and systems to promote shared prosperity. Training for Employment, Productivity and Social Inclusion. International Labour Conference, Geneva.

Lynch, L. M., Black, S.E., 1998. Beyond the incidence of training: Evidence from a national employers survey. Industrial and Labor Relations Review 51, 64-81.

Manning, A., 2003. Monopsony in motion: Imperfect competition in labor markets. Princeton University Press, Princeton.

Manning, A. 2011. Imperfect competition in the labor market, in: Card, D., Ashenfelter, O. (Eds.), Handbook of Labor Economics vol 4b, Elsevier Science B.V., pp. 973-1041.

Manning, A., 2021. Monopsony in the labor market: A review. Industrial and Labor Relations Review 74, 3-26.

McCall, B., Smith, J.A., Wunsch, C., 2016. Government-sponsored vocational education for Adults, in: Hanushek, E.A., Machin, S., Woessmann, L. (Eds.), Handbook of the Economics of Education vol 5, Elsevier Science B.V., pp. 479-652.

McConnell, S., Schochet, P., Rotz, D., Fortson, K., Burkander, P., Mastri, A., 2021. The effects of employment counseling on labor market outcomes for adults and dislocated workers: Evidence from a nationally representative experiment. Journal of Policy Analysis and Management 40, 1249-1287.

Meyer, B., Mok, W., Sullivan, J. 2015. Household surveys in crisis. Journal of Economic Perspectives 29, 199-226.

Miller, C., Hedlam, C., Mano, M., Cullinan, D., 2020. Increasing community college graduation rates with a proven model: Three year results from the Accelerated Study in Associates Program (ASAP) Ohio Demonstration. Manpower Demonstration Research Corporation, New York.

Mincer, J., 1958. Investments in human capital and personal income distribution. Journal of Political Economics 66, 281-392.

Mincer, J., 1962. On-the-job training: Costs, returns, and some implications. Journal of Political Economy 70, 50-79

Mincer, J., 1974. Schooling, experience and earnings. Columbia University Press, New York.

Mincer, J., 1980. Union effects: wages, turnover, and job training. NBER Working Paper No. 0808.

Mincer, J., 1988. Job training, wage growth and labor turnover. NBER Working Paper No. 2090.

Mincer, J., Jovanovic, B., 1981. Labor mobility and wages, in: Rosen, S. (Ed.) Studies in labor markets, University of Chicago Press, Chicago, pp. 21-64.

Mincer, J., Leighton, L., 1980. Effect of minimum wages on human capital formation. NBER Working Paper No. 441.

Moen, E., Rosén, A., 2004. Does poaching distort training? Review of Economic Studies 71, 1143-1162.

Moffitt, R., 1991. Program evaluation with nonexperimental data. Evaluation Review. 15(3), 291-314.

Mohrenweiser, J., 2016. Recruitment and apprenticeship training. Industrielle Beziehungen 23, 6-24.

Mohrenweiser, J., Zwick, T., Backes-Gellner, U., 2019. Poaching and firm-sponsored training. British Journal of Industrial Relations 57, 143-81.

Morgan, P., Manning, R., 1985. Optimal search. Econometrica 53, 923-944.

Mortensen, D.T., 1978. Specific capital and labor Turnover. Bell Journal of Economics 9, 572-86.

Muehler, G., Beckmann, M., Schauenberg, B., 2007. The returns to continuous training in Germany: New evidence from propensity score matching estimators. Review of Managerial Science 1, 209-235.

Muehlemann, W., Ryan, P., Wolter, S.C., 2013. Monopsony power, pay structure, and training. Industrial and Labor Relations Review 66, 1097-114.

Murphy, K.M., Welch, F., 1992. The structure of wages. Quarterly Journal of Economics 107, 285-326.

Mueller, N., 2014. Does CVT of firms in Germany suffer from poaching? Empirical Research in Vocational Education and Training 6, 1-26.

Myerson, R.B., Satterthwaite, M.A., 1983. Efficient mechanisms for bilateral trading. Journal of Economic Theory 29, 265-81.

Nagypál, É., 2005. On the extent of job-to-job transitions. mimeo.

Neal, D., 1995. Industry-specific human capital: Evidence from displaced workers. Journal of Labor Economics 13, 653-677.

Neal. D., Johnson, W.R., 1996. The role of premarket factors in black-white wage differences. Journal of Political Economy 104, 869-895.

Neumark, D., Wascher, W., 2001. Minimum wages and training revisited. Journal of Labor Economics 19, 563-95.

Ohashi, I., 1983. Wage profiles, layoffs and specific training. International Economic Review 24, 169-81.

Oi, W.Y., 1962. Labor as a quasi-fixed factor. Journal of Political Economy 70, pp. 538-55

Orr, L, Bloom, H., Bell, S., Doolittle, F., Lin, W., Cave, G., 1996. Does training for the disadvantaged work? Evidence from the national JTPA Study. Urban Institute Press, Washington DC.

OECD, 1998. Education at a Glance - OECD Indicators. OECD, Paris.

Parent, D., 1999. Wages and mobility: The impact of employer-provided training. Journal of Labor Economics 17, 298-317.

Parsons, D.O., 1972. Specific human capital: An application to quite rates and layoff rates. Journal of Political Economy 80, 1120-43.

Peck, L., Litwok, D., Walton, D., 2022. Health Profession Opportunity Grants (HPOG 1.0) impact study: Six-year impacts report. Abt Associates, Rockville.

Peracchi, F., 2002. The European Household Community Panel: A review. Empirical Economics 27, 63-90.

Picchio, M., van Ours, J.C., 2011. Market imperfections and firm-sponsored training. Labour Economics 18, 712-722.

Pigou, A., 1912. Wealth and welfare. Macmillan, London.

Pischke, J.S., 2001. Continuous training in Germany. Journal of Population Economics 14, 523-548

Polachek, S. W., 1981. Occupational self-selection: a human capital approach to sex differences in occupational structure. Review of Economics and Statistics 63, 60-69.

Polachek, S.W., 1985a. Occupation segregation: a defence of human capital predictions. Journal of Human Resources 20, 437-440.

Polachek, S.W., 1985b. Occupational segregation: Reply to England. Journal of Human Resources 20, 444.

Rzepkaa, S., Tamm, M., 2016. Local employer competition and training workers. Applied Economics 45, 3307–3321.

Roder, A., Elliot, M., 2018. Escalating gains: The elements of project QUEST's success. Economic Mobility Corporation, New York.

Rosen, S., 1972a. Learning by experience as joint production. Quarterly Journal of Economics 86, 366-82.

Rosen, S., 1972b. Learning and experience in the labor Market. Journal of Human Resources 7, 362-42.

Roser, M., Ortiz-Ospina, E., 2013. Income inequality. Our World in Data.

Rosholm, M., Nielsen, H.S., Dabalen, A., 2007. Evaluation of training in African enterprises. Journal of Development Economics 84, 310-329.

Ruhose, J., Thomsen, S.L., Weilage, I., 2019. The benefit of adult learning: Work-related training, social capital and earnings. Economics of Education Review 72, 166-186.

Sacerdote, B. 2017. Fifty years of growth in American consumption, income, and wages. NBER Working Paper No. 23292.

Sauermann, J., Stenberg, A., 2020. Assessing selection bias in non-experimental estimates of the returns to workplace training. IZA Discussion Paper No. 13789.

Schultz, T.W., 1961. Investment in human capital. American Economic Review 51, 1-17.

Schumann, M., 2017. The effects of minimum wages on firm-financed apprenticeship training. Labour Economics 47, 163-181.

Schwerdt, G., Messer, D., Woessmann, L., Wolter, S.C., 2012. The impact of an adult education voucher program: Evidence from a randomized field experiment. Journal of Public Economics 96, 569-583.

Sheehan, K.B. and S. McMillan, 1999. Response variation in e-mail surveys: An exploration. Journal of Advertising Research 39, 45-54

Sianesi B. 2004. An evaluation of the Swedish system of active labor market programs in the 1990s. Review of Economics and Statistics 86(1), 133–155.

Smith, A., 1776. An inquiry into the nature and causes of the wealth of nations. W. Strahan and T. Cadell, London.

Smith, J.A., Whalley, A., 2006. How well do we measure public job training? mimeo.

Spengler, J.J., 1977. Adam Smith on Human Capital. The American Economic Review 67, 32-36.

Starr, E., 2019. Consider this: Training, wages, and the enforceability of covenants not to compete. Industrial and Labor Relations Review 72, 783-817.

Starr, E., Prescott, J.J., Bishara, N.D., 2020. The behavioral effects of (unenforceable) contracts. Journal of Law, Economics, and Organization 36, 633-87.

Starr, E., Prescott, J.J., Bishara, N.D., 2021. Noncompete agreements in the US labor force. Journal of Law and Economics 64, 53-84.

Stevens, M., 1994. A theoretical model of on-the-job training with imperfect competition. Oxford Economic Papers 46, 537-62.

Stevens, M., 2001. Should firms be required to pay for vocational training? Economic Journal 111, 485-505.

Stith, S. 2018. Organizational learning-by-doing in liver transplantation. International Journal of Health Economics and Management 18, 25–45.

Stockinger, B., Zwick, T., 2017. Apprentice poaching in regional labor markets. ZEW

Tamm, M., 2018. Training and changes in job tasks. Economic of Education Review 67, 137-147.

Topel, R., 1986. Job mobility, search, and earnings growth: A reinterpretation of human capital earnings functions. Research in Labor Economics 8, 199-233.

Topel, R., 1990. Specific capital and unemployment: Measuring the costs and consequences of job loss. Carnegie-Rochester Conference Series on Public Policy 33, 181-214.

Topel, R., 1991. Specific capital, mobility and wages: Wages rise with job seniority. Journal of Political Economy 99, 145-75.

Turcotte, J., Léonard, A., Montmarquette, C., 2003. New evidence on the determinants of training in Canadian business locations. Statistics Canada. The Evolving Workplace Series No. 71.

VanderWeele, T., Hernan, M. 2013. Causal inference under multiple versions of treatment. Journal of Causal Inference 1, 1–20.

Wallace, G., Haveman, R., 2007. The implications of differences between employer and worker employment/earnings reports for policy evaluation. Journal of Policy Analysis and Management 26, 737-753.

Wolter, S. C., Ryan, P., 2011. Apprenticeship, in: Hanushek, E.A, Machin, S., Woessmann, L. (Eds.), Handbook of the Economics of Education vol 3, Elsevier Science B.V., pp. 521-576.

Young, S., 2021. Noncompete clauses, job mobility, and job quality: Evidence from a low-earning noncompete ban in Austria. SSRN Working Paper No. 3811459.

9 Tables

TABLE 1: TRAINING QUESTIONS IN REFERENCED SURVEYS

USA

National Longitudinal Survey of Youth 1979 Cohort (NLSY-79)

"(Besides the training we've already talked about) Since (DATE of LAST INTERVIEW), have you received training from any (other) source, such as the kinds of places listed on this card? For example, training in a business college, nurse program, an apprenticeship program, a vocational-technical institute, or any of these other kinds of source?"
(The questionnaire contains detailed questions for up to three training spells since last interview.)

National Longitudinal Survey of Youth 1997 Cohort (NLSY-97)

"In addition to the programs which we have already talked about, since [date of last interview], have you attended any schooling, course or training programs designed to help people find a job, improve their skills, or learn a new job? Some sources of occupational training programs include business schools, cosmetology schools, nursing courses, apprenticeships, vocational or technical institutes or schools, correspondence courses, company or military trainings, employer training programs, and night schools."
(The questionnaire contains detailed questions about all training spells since last interview.)

Employment Opportunity Pilot Program 1982 (EOPP)
and *Small Business Administration* 1992 (SBA)

"During the first three months of work, what was the total number of hours spent on formal training, such as self-paced learning programs or training done by specially trained personnel?"

"During the first three months of work, what was the total number of hours management and line supervisors spent away from other activities giving informal individualized training or extra supervision?"

"During the first three months of work, what was the total number of hours co-workers who are not supervisors spent away from their normal work giving informal individualized training or extra supervision?"

"During the first three months of work, how many total hours does the average new employee spend in training activity in which he or she is watching other people rather than doing it himself or herself?"

Panel Study of Income Dynamics 1976, 1978, and 1985 (PSID)

"On a job like yours, how long would it take the average new person to become fully trained and qualified?"

Panel Study of Income Dynamics 1993 (PSID)

"Suppose someone had the experience and education needed to start working at a job like yours. From that point, how long would it take them to become fully trained and qualified (to do a job like yours)?

Current Population Survey 1991*, Job Training Supplement* (CPS)

"Did you need specific skills or training to obtain your current (last) job?"

"Since you obtained your present job did you take any training to improve your skills?"

(The questionnaire contains follow-up questions on type, duration, supplier, and financing of training taken.)

Survey of Employer-Provided Training 1995 (SEPT)

Employee part: "In the last 12 months, have you received formal training from your current employer in any of the following areas?" (A list of 11 different types of training is presented to the employee from 'management training' to 'basic skills training'.)"

"Fill out a worksheet at the end of each day, while the day's activities are still fresh in your mind. Do this every day for one week even if there were days with no training activities." (The employee is presented with information about what constitutes training according to BLS. Options provided for each of the following:

1) Did you learn a skill or were you given information today to help you do your job better?
2) Who or what helped you learn this skill or information?
3) How did you learn the skill or information?
4) What type of skill or information did you learn?
5) How long did you spend learning this skill or information today?)

Employer part: "In the last 12 months, in which of the following areas did your establishment provide or finance training for employees?" (The same list supplied as for employees plus an additional three – employee health and wellness, orientation training, and awareness training.) "Please complete one worksheet for every formal training activity that takes place during your two-week reporting period." (The sheet includes duration of activity and number of employees attending.)

CAN *Access and Support to Education and Training Survey* (ASETS)

"Between July 2007 and June 2008, (including the program or programs we have already discussed) how many educational programs in total have you been enrolled in? By program we mean a series of courses taken towards a diploma, certificate, degree or license, which normally takes more than 3 months to complete."

"Between July 2007 and June 2008, (other than the program or programs just mentioned) how many courses, workshops or seminars did you take?"
(The questionnaire contains detailed questions about the most recent formal program taken and up to a maximum of ten of the informal.)

AUS *Household, Income and Labour Dynamics in Australia* (HILDA)

"During the past 12 months, have you taken part in any education or training schemes, as part of your_employment?"
(The questionnaire contains detailed information about all training episodes aggregated.)

UK *National Child Development Survey* 1991 (NCDS)

"Since March 1981 have you been on any courses that were meant to lead to qualifications?"

"Since March 1981 have you been on any training courses designed to help you develop skills that you might use in a job?"
(The questionnaire contains detailed questions about the two courses leading to the highest qualification and the three most recent training courses over the ten-year period.)

British Household Panel Survey (BHPS*)*

"(Apart from the full-time education you have already told me about) Have you taken part in any other training scheme or courses at all since September 1st [last year] or completed a course of training which led to a qualification? Please include part-time college or university courses, evening classes, training provided by an employer either on or off the job, government training schemes, Open University courses, correspondence courses and work experience schemes."
(Respondents are explicitly told to exclude leisure courses. The questionnaire contains detailed questions about the three "longest" training spells.)

GER *German Socioeconomic Panel* (SOEP*)*

"If you want to continue your professional education, there are various options. Think about the last three years. During this time, have you [...] taken part in job-related training programs or courses, including those that are still running?"
(The questionnaire contains detailed questions about the "most important" course.)

Further Training as a Part of Lifelong Learning (WeLL)

"Please tell me whether you have taken part in the following further training opportunities in the period from January 2006 to the present. We are only interested in job-related training and further education. Have you taken part in...

a. Seminars or training courses that were <u>not</u> carried out by your company but by an external organization?
b. Seminars or training courses carried out by your company?"

(The questionnaire collects detailed information for up to three most recent events over a two-year period. Interviews done from October 2007 to January 2008. Additional questions on OJT and workshops included as well.)

OECD *The International Adult Literacy Survey* (IALS)

"During the past 12 months, that is, since [month, year], did you receive any training or education including courses, private lessons, correspondence courses, workshops, on-the-job-training, apprenticeship training, arts, crafts, recreation courses or any other training or education?"
(The questionnaire contains detailed information about the three most recent activities.)

The Programme for the International Assessment of Adult Competencies (PIAAC)

"During the last 12 months, that is since [month, year], have you studied for any formal qualification, either full-time or part-time?"
"During the last 12 months, have you attended

a. any organized sessions for on-the-job training or training by supervisors or co-workers?
b. Seminars or workshops?

c. Courses or private lessons, not already reported?"

"Still talking about your current job: If applying today, what would be the usual qualifications, if any, that someone would need to GET this type of job?" (A list of 14 different educational levels is presented)

"Thinking about whether this qualification is necessary for doing your job satisfactorily, which of the follow statements would be most true?" (Three options are presented: lower would be sufficient, the level is necessary, and a higher level needed.)

"Suppose that someone with this level of qualification were applying today, how much related work experience would they need to GET this job?" (A list of six options presented. From "None" to "3 years or more".)

(The questionnaire collects detailed information about the most recent formal and informal activity.)

EU

European Community Household Panel (ECHP)

"Have you at any time since January [last year] attended a course in general or higher education?"

"Have you at any time since January [last year] been in vocational education or training, including any part-time or short courses?"

(The questionnaire collects some information about the most recent general and most recent vocational training activities including level.)

Adult Education Survey (AES)

"During the last 12 months, that is since [month, year] have you been a student or apprentice in formal education or training?"

"During the last 12 months have you participated in

a. Courses at the workplace or in your free time?
b. Workshops or seminars at the workplace or in your free time?
c. Guided on-the-job training, which means planned periods of education, instructions or training directly at the workplace, organised by the employer with the aid of a designated teacher/instructor?
d. Private lessons with the aid of a teacher or tutor for whom this is a paid activity?"

(The questionnaire contains detailed information about the most recent formal activity and two randomly selected informal activities out of at most ten.)

Continuing Vocational Training Survey (CVTS)

"In 2015, did your enterprise provide internal or external CVT courses?"

"In 2015, did your enterprise provide any of the following other forms of CVT?"

(The questionnaire contains a list of six different types of training along with an assessment of the fraction of workers participating grouped into brackets less than 10%, from 10% to 50%, and 50% or more.)

Note: The *NLSY-79* (wording from round 5, 1983) and *NLSY-97* (wording from round 3, 1999) questions are lifted from the webpage of the U.S. Bureau of Labor Statistics. The *EOPP* and *SBA* questions are lifted from questionnaires in the authors' possession. The formal training question in SBA was separated into one on onsite and one on offsite formal training. The *PSID* questions are lifted from the ISR webpage at University of Michigan. The *CPS* questions were lifted from www2.census.gov. The *SEPT* is taken from Lerman et al. (1999). The *ASETS* is taken from the webpage of Statistics Canada. The *HILDA* is taken from the webpage of the Melbourne Institute, The University of Melbourne. The *NCDS* is taken from the webpage of the Centre for Longitudinal Studies at UCL. The *BHPS* is taken from the webpage of the Institute for Social and Economic Research at

University of Essex. The *SOEP* is taken from the webpage of DIW. The *WeLL* is taken from the webpage of IAB. The two surveys were translated into English by the authors. The *PIAAC* is taken from the webpage of the OECD. The EU surveys information is taken from the webpage of Eurostat.

Table 2: 40 Years of Measuring Training

Survey Name	Countries Covered	Period Surveyed	Population and Sample Size	Training Measures [1)]	Incidence rates	Time spent in training among participants	Example of participation study
Household surveys							
National Longitudinal Survey of Youth 1979 (NLSY79)	USA	1979-2018	12,686 men and women born 1957-64 (6,878 by 2018)	Formal and informal training	0.17 for formal and 0.40 for informal (previous 12 months)	50-60 hours (previous 12 months)	Loewenstein and Spletzer (1999)
Adult Education and Training Survey (AETS)	CA	1983, 1985, 1989, 1991, 1993, 1997, 2002	33,000 respondents aged 16 and above (1997)	Training programs leading to a formal certificaton and training courses that do not.	0.07 for training programs and 0.22 training courses (previous 12 months). Rates for 25–64-year-old non-full-time population.	363 hours for programs and 39 hours courses	Hui and Smith (2002)
European Community Household Panel (ECHP)	AT, BE, DE, DK, ES, FR, GB, GR, IT, IE, LU, NL, PT, SE	1994-2001	Around 80,000 households in the 14 member states.	Vocational education or training, incl. part-time and short courses (no distinction possible).	0.41 in UK and DK to 0.06 in IE and IT. Rates for the 25–54-year-old working population, since Jan last year.		Arulampalam, et al. (2004)
International Adult Literacy Survey (IALS)	CA, DE, IE, NL, PL, SW, CH, USA (1st round); AUS, GB, BE, NZ (2nd); Chile, FI, NO, CZ, HU, SL, DK, IT, CH (3rd)	1994 (1st round), 1996 (2nd round), 1998 (3rd round)	Between 2,000-4,500 per nation	Any training or education participation. Three most recent episodes can be separated into job-related or not and training towards a degree or not.	From as high as 0.46 in the US and 0.52 in the UK to 0.17 in PO and 0.20 in BE. Rates for job-related education and training for 25–64-year-old working population, previous 12 months [3)]	131 (BE)-271 (IE) hours (previous 12 months)	Leuven and Oosterbeek (1999)

Adult Education Survey	BE, BG, CZ, DK, DE EE, IE, EL, ES, FR, IT, CY, LV, LT, LU, HU, MT, NL, AT, PL, PT, RO, SI, SK, FI, SE, UK, NO, CH, RS	2007 (pilot), 2011, 2016, 2022	240,000 respondents, ages 25 to 64.	Formal and informal training	Formal: from as low as 0.02 in BG, GE, GR, IT, SL and RO to 0.16 in the UK. Informal: From as low as 0.12 in Greece and 0.29 in the UK to as 0.66 in NL and 0.75 in SW. Rates for employed respondents, previous 12 months.[2)]		
The Programme for the International Assesment of Adult Competencies (PIAAC)	AUS, AT, BE, CAN, CZ, DK, ES, FI, FR, GE, IR, IT, JA, KO, NL, NO, PO, RUS, SL, ES, SW, UK, US (round 1); CH, GR, IDN, ISR, LT, NZ, SG, SE, TUR (round 2); ECU, HU, KAZ, MEX, PERU, US (Round 3).	2011 (round 1), 2014-15 (round 2), 2017 (round 3)	250,000, ages 16 to 65.	Formal and informal training	Formal: From as low as 0.02 and 0.03 in JAP and GER to 0.19 for TUR. Informal: from as low as 0.16 and 0.19 in KAZ and RUS to 0.70 in NL, SW, DK, FI, and NZ. Rates for the private-sector employed respondents, previous 12 months.	85 hours of job-related informal training and 28 hours of other types of informal training for the US sample of private-sector employees per year.	This chapter

Employer surveys							
Employment Opportunity Pilot Project 1982 (EOPP)	USA	1982	1,916 firms across 23 sites. Oversampling of low-income workers.	Formal and informal training of most recently hired worker.	0.15 for formal and 0.96 for informal (first three months of employment).	70-80 hours (during first three months of employment).	Barron et al. (1997b)
Small Business Administration (SBA)	USA	1992	1,288 firms. Oversampling of large firms.	Formal and informal training of most recently hired worker	0.20 for formal and 0.98 for informal (first three months of employment)	65-70 hours (during first three months of employment)	Barron et al. (1997b)
Educational Quality of the Workforce National Employers Survey (EQW-NES)	USA	1994	2,945 firms with more than 20 workers	Formal training and informal training.	0.40 formal training (no timeframe given).	N.A. Costs: 0.05 of total labor costs.	Lynch and Black (1998)
Ekos Workplace Training Survey (WTS)	CA	1995	2,584 establishments with at least one employee	Formal and informal training supplied by firm over past year	0.42 formal training and 0.70 informal training	N.A.	Betcherman et al. (1997)
The Continuing Vocational Training in Enterprises Survey (CVTS)	The 28 EU member states and NO	1995, 2000, 2005, 2010, 2015, 2020	Firms (≥ 10 workers) except agriculture, forestry and fishing, public administration and defense, compulsory social security, education, human health and social work activities	Formal and informal training.	Formal: 0.38 for the 28 EU countries but as low as 0.16 in GR and as high as 0.61 in CZ. Informal: 0.20 on average for the 28 EU countries (previous 12 months) [2)]	N.A. Costs: 0.023 of total labor costs (from 1.0 in BG to 3.0 in DK)	Bassanini et al. (2007)

Matched employer-employee surveys

Multi-City Study of Urban Inequality (MCS)	USA	1992-1994	Roughly 2,200 establishments across metropolitan areas of Atlanta, Boston, Detroit, and Los Angeles	Formal and informal training received by most recently hired worker	0.44 for formal training and 0.91 for informal (no information regarding time period)	Between 52 hours (hispanics) and 116 (white men) in formal and 72 (black females) and 99 (white men) in informal.	Holzer and Reaser (1999)
Survey of Employer-Provided Training (SEPT)	USA	1995	1,062 establishments with more than 50 workers and 949 of their workers	Formal and informal training classified using a Bureau of Labor Statistics algorithm.	0.70 formal training over previous 12 months (worker response)	unconditional estimate: 0.04 of working hours over a year spent in training and $3,050 (measured in 2022-$)	Frazis et al. (2000)
Workplace and Employee Survey (WES)	CA	1999-2006	6,322 establishments and 23,540 workers (1999)	Classroom training with predetermined format, objective, and content. Informal like orientation, supervision, team building, and health and safety.	0.20 classroom training and 0.30 informal (employers) and 0.37 classroom and 0.30 informal (workers)	N.A.	Turcotte et al. (2003)

NOTE: [1] Further description of the training measures: *NLSY* formal covers apprenticeships, business and vocational technical schools, on-the-job training (OJT), inside and outside seminars, and other training. Informal training covers classes, seminars, instructions from supervisor and coworkers, and self-study using manuals, workbooks, or computers. *AES* and *PIAAC* formal covers education and training leading to a qualification. Informal training covers courses, workshops, seminars, guided OJT, and lessons by paid teacher. *EOPP* and *SBA* formal covers self-paced learning programs and training done by specially trained personnel. Informal covers individualized training and extra supervision by either management, supervisor, or co-workers. *EQW-NES* formal covers structured training with predefined objectives such as seminars, lectures, workshops, audio-visual presentations, apprenticeships, and structured OJT. Informal covers learning by watching others, and one-to-one instruction on job tasks. *WTS* formal covers classroom instruction, scheduled and structured OJT, apprenticeships, and courses at formal educational institutes. Informal is residual training. *CVTS* formal covers continuing vocational training managed internally or externally. Informal training covers planned periods of training, instruction, or practical experience, planned learning through job rotations, learning circles, self-learning through open or distance learning, or conferences, workshops, lectures, and seminars. *MCS* formal covers training by a professional instructor. Informal training by either co-worker or supervisor. [2] Numbers taken from Goglio and Meroni (2014) [3] Number taken from Education at a Glance, OECD (1998), Tables C5.1 and C5.2

TABLE 3: INCIDENDE AND HOURS OF TRAINING ACROSS SELECTED COUNTRIES

	United States	United Kingdom	Germany	OECD
Formal Training, pct				
Overall	12.6	15.6	3.1	9.1
Labor market status at time of training:				
... Employed	11.7	14.7	2.1	8.3
... Not employed	0.9	0.8	1	0.8
Financing of training:				
... (Co-)financed by employer	6.6	11.3	1.7	5.2
... No employer financing	6.0	4.3	1.4	3.9
Main reason for participation:				
... Job-related	9.6	13.3	2.6	6.9
... Other reasons	3.0	2.3	0.5	2.2
Informal training, pct				
Overall	65.7	63.0	59.9	54.0
Training activity type:				
... Open or distance courses	17.9	8.1	7.3	10.1
... Organized on-the-job training	53.3	49.1	48.3	40.4
... Workshops or seminars	36.1	33.7	29.9	24.8
... Private lessons	8.6	7.5	10.4	9.6
Labor market status at time of training:				
... Employed	65.4	62.5	59.2	53.0
... Not employed	0.3	0.5	0.7	1.0
Financing of training:				
... (Co-)financed by employer	62.1	61.1	56.6	50.2
... No employer financing	3.6	1.9	3.3	3.8
Main reason for participation:				
... Job-related	61.2	59.3	55.5	49.2
... Other reasons	4.5	3.7	4.4	4.8
Hours of informal training conditional on participation:				
... Job-related	85	65	69	71
	(242)	(122)	(167)	(164)
... Other training	28	13	10	30
	(178)	(101)	(58)	(132)
Casual training, pct	93.6	88.4	91.6	88.3
Learning by doing, pct	94.2	92.0	96.7	91.7

Sample size	1,764	1,998	1,512	39,870

Note: Training measures are taken from the PIAAC sample of 25–59-year-olds working full-time in the private sector. Standard deviations appear in parentheses. We define employer co-financed training as that where the employer pays (in full or in part) the direct costs of training and/or where the training took place (mostly or entirely) during working hours. Sampling weights are used to make the data representative of the target populations. OECD column is (unweighted) country averages of the available OECD member states.

Table 4: Effects of Worker, Job, and Establishment Characteristics on the Probability of Formal Training Participation, US Sample of PIAAC

	Participants	Non-participants	Std. diff.	Coeff.		P-value of F-test
Worker characteristics:						<0.0001
Male	0.486	0.580	-16.5%	-0.015		
Age indicators:						<0.0001
... 25≤ age ≤ 34	0.551	0.279	52.2%	0.099	***	
... 35 ≤ age ≤ 44	0.245	0.291	-8.9%	(ref)		
... 45 ≤ age ≤ 54	0.133	0.313	-37.6%	-0.062	***	
... 55 ≤ age ≤ 59	0.071	0.117	-12.8%	-0.034	*	
Married	0.622	0.714	-18.1%	-0.030		
Native	0.846	0.806	8.9%	0.013		
Number of children:						0.227
... 0	0.397	0.260	26.7%	0.045		
... 1	0.140	0.172	-7.9%	(ref)		
... 2	0.204	0.289	-17.5%	0.025		
... 3+	0.258	0.279	-3.9%	0.060	**	
Child below the age of six in household	0.181	0.145	8.8%	-0.006		
Educational attainment:						0.712
... Primary and lower secondary	0.042	0.091	-16.5%	-0.012		
... Upper secondary	0.408	0.479	-12.4%	(ref)		
... Tertiary	0.550	0.431	21.1%	0.018		
School dropout (any level):	0.223	0.285	-12.4%	-0.044	**	
Job characteristics:						0.005
Index of use of skills at work:						0.005
... ICT	-0.017	0.002	-3.6%	-0.053	**	
... Numeracy	0.093	-0.013	17.7%	0.040	*	
... Reading	0.089	-0.013	16.2%	0.013		
... Writing	0.131	-0.019	23.0%	0.059	**	
Managing other employees	0.375	0.412	-6.5%	-0.040	**	
Occupation (ISCO-based):						0.008
... Skilled	0.632	0.520	19.9%	0.022		
... Semi-skilled, white-collar	0.243	0.219	5.1%	(ref)		
... Semi-skilled, blue-collar	0.086	0.205	-29.6%	-0.069	**	
... Elementary	0.038	0.055	-7.1%	-0.032		
Hourly wage rate, in deciles						0.576
... Lowest decile	0.034	0.037	-1.2%	0.017		
... 2nd decile	0.073	0.062	3.8%	0.034		
... 3rd decile	0.062	0.084	-7.7%	-0.031		
... 4th decile	0.118	0.093	7.3%	0.031		

... 5th decile	0.108	0.096	3.7%	(ref)		
... 6th decile	0.120	0.110	3.0%	-0.020		
... 7th decile	0.121	0.102	5.3%	-0.016		
... 8th decile	0.110	0.111	-0.3%	-0.024		
... 9th decile	0.123	0.111	3.2%	-0.046		
... Highest decile	0.080	0.121	-11.8%	-0.068	*	
Establishment characteristics:						0.007
Establishment size:						0.202
... 1 to 10 people	0.172	0.179	-1.6%	-0.022		
... 11 to 50 people	0.239	0.285	-9.3%	-0.051	**	
... 51 to 250 people	0.285	0.240	9.1%	(ref)		
... 251 to 1,000 people	0.143	0.173	-7.2%	-0.039		
... More than 1,000 people	0.156	0.122	8.7%	-0.005		
Over the last 12 months, the establishment has						0.135
... Increased the number of workers	0.315	0.278	7.2%	0.028		
... Stayed more or less the same	0.446	0.526	-19.9%	(ref)		
... Decreased the number of workers	0.225	0.188	7.9%	0.042	*	
Establishment part of larger firm	0.624	0.610	2.5%	0.005		
Industry (ISIC):						0.011
... Agriculture, forestry, and fishing (A)	0.010	0.008	2.4%	0.016		
... Mfg + industrial (B,C,D,E)	0.183	0.206	-5.0%	-0.074		
... Construction (F)	0.070	0.052	6.5%	(ref)		
... Wholesale, retail, food (G,H,I)	0.123	0.233	-26.5%	-0.142	***	
... Information and communication (J)	0.058	0.071	-4.6%	-0.096	*	
... Financial and insurance activities (K)	0.115	0.076	11.2%	-0.048		
... Real estate activities (L)	0.009	0.014	-4.9%	-0.139	*	
... Pro, sci, and admin services (M, N)	0.092	0.110	-4.9%	-0.122	**	
... Pub adm, edu, and health (O,P,Q)	0.294	0.183	23.6%	-0.067		
... Other service activities (R,S,T,U)	0.045	0.046	-0.6%	-0.107		
R^2				0.109		
Dependent variable mean				0.126		
Unweighted N	256	1,508				

Note: Formal training participation is measured using question *B_ Q04a* in the *PIACC* questionnaire. Std diff: standardized difference (difference in sample means of the two types of respondents divided by the square root of the mean of the empirical variances in the two samples). Coeff: estimated regressions coefficients. Regression based on the US PIAAC sample of 25–59-year-olds working full-time in the private sector. Model also includes an intercept as well as indicators for item non-response. *** = p<0.01; ** = p < 0.05; * = p < 0.10. The F-tests are tests of the joint null that the set of population coefficients of all included variables in a given set equal zero. The partial R^2 of worker, job, and establishment characteristics equal 0.045, 0.028, and 0.021. Weighted means and regression.

Table 5: Effects of Worker, Job, and Establishment Characteristics on the Probability of Informal Training Participation, US Sample of PIAAC

	Participants	Non-participants	Std diff	Coeff.	P-value of F-test
<u>Worker characteristics:</u>					0.044
Male	0.550	0.602	-9.4%	-0.002	
Age indicators:					0.425
... 25≤ age ≤ 34	0.315	0.309	1.4%	0.019	
... 35 ≤ age ≤ 44	0.285	0.285	0.0%	(ref)	
... 45 ≤ age ≤ 54	0.290	0.290	0.0%	-0.006	
... 55 ≤ age ≤ 59	0.109	0.116	-1.7%	-0.057	
Married	0.722	0.666	11.2%	0.037	
Native	0.844	0.748	20.4%	0.015	
Number of children:					0.453
... 0	0.294	0.246	9.9%	0.038	
... 1	0.169	0.165	0.9%	(ref)	
... 2	0.295	0.246	9.7%	0.036	
... 3+	0.242	0.343	-19.0%	-0.008	
Child below the age of six in household	0.141	0.165	-5.9%	-0.032	
Educational attainment:					0.009
... Primary and lower secondary	0.041	0.167	-35.2%	-0.011	
... Upper secondary	0.409	0.586	-31.5%	(ref)	
... Tertiary	0.550	0.246	57.7%	0.096***	
School dropout (any level):	0.292	0.248	8.8%	0.072***	
<u>Job characteristics:</u>					<0.0001
Index of use of skills at work:					<0.0001
... ICT	0.039	-0.075	22.3%	0.001	
... Numeracy	0.040	-0.077	21.6%	-0.003	
... Reading	0.118	-0.227	59.3%	0.109***	
... Writing	0.070	-0.135	36.1%	0.061**	
Managing other employees	0.443	0.337	19.0%	-0.023	
Occupation (ISCO-based):					0.441
... Skilled	0.646	0.321	60.4%	0.059	
... Semi-skilled, white-collar	0.190	0.282	-19.2%	(ref)	
... Semi-skilled, blue-collar	0.137	0.292	-33.8%	0.020	
... Elementary	0.026	0.104	-28.1%	-0.007	
Hourly wage rate, in deciles					0.216
... Lowest decile	0.020	0.069	-20.8%	-0.044	
... 2nd decile	0.043	0.103	-19.8%	0.009	
... 3rd decile	0.063	0.115	-16.5%	-0.010	
... 4th decile	0.081	0.124	-12.4%	0.032	

... 5th decile	0.086	0.119	-9.6%	(ref)		
... 6th decile	0.109	0.115	-1.6%	0.003		
... 7th decile	0.108	0.097	2.9%	-0.001		
... 8th decile	0.135	0.066	20.7%	0.110	**	
... 9th decile	0.144	0.053	27.7%	0.088	*	
... Highest decile	0.148	0.054	26.7%	0.090	*	
Establishment characteristics:						0.001
Establishment size:						0.541
... 1 to 10 people	0.159	0.216	-12.7%	-0.005		
... 11 to 50 people	0.264	0.309	-8.9%	-0.026		
... 51 to 250 people	0.248	0.241	1.3%	(ref)		
... 251 to 1,000 people	0.172	0.163	2.0%	-0.047		
... More than 1,000 people	0.157	0.068	24.6%	0.018		
Over the last 12 months, the establishment has						0.710
... Increased the number of workers	0.303	0.245	11.4%	0.003		
... Stayed more or less the same	0.496	0.552	-9.9%	(ref)		
... Decreased the number of workers	0.189	0.199	-2.3%	-0.024		
Establishment part of larger firm	0.665	0.508	28.5%	0.127	***	
Industry (ISIC):						0.131
... Agriculture, forestry, and fishing (A)	0.003	0.018	-12.9%	-0.141		
... Mfg + industrial (B,C,D,E)	0.187	0.235	-10.6%	-0.072		
... Construction (F)	0.046	0.071	-9.5%	(ref)		
... Wholesale, retail, food (G,H,I)	0.178	0.296	-25.0%	-0.109	*	
... Information and communication (J)	0.084	0.040	16.1%	-0.059		
... Financial and insurance activities (K)	0.099	0.045	17.8%	-0.011		
... Real estate activities (L)	0.013	0.014	-0.2%	-0.110		
... Pro, sci, and admin services (M, N)	0.110	0.104	1.5%	-0.078		
... Pub adm, edu, and health (O,P,Q)	0.234	0.126	24.6%	0.016		
... Other service activities (R,S,T,U)	0.045	0.048	-1.2%	-0.042		
R^2				0.245		
Dependent variable mean				0.657		
Unweighted N	1,171	593				

Note: Informal training participation is measured using question *B_ Q13* in the *PIACC* questionnaire. Std diff: standardized difference (difference in sample means of the two types of respondents divided by the square root of the mean of the empirical variances in the two samples). Coeff: estimated regressions coefficients. Regression based on the US 2012-sample of 25–59-year-olds working full-time in the private sector. Model also includes an intercept as well as indicators for item non-response. *** = p<0.01; ** = p < 0.05; * = p < 0.10. The F-tests are tests of the joint null that the set of population coefficients of all included variables in a given set equal zero. The partial R^2 of worker, job, and establishment characteristics equal 0.014; 0.080, and 0.025. Weighted means and regression.

Table 6: Effects of Worker, Job, and Establishment Characteristics on the Probability of Formal Training Participation, Full PIAAC Sample

	Average $\hat{\beta}_c$	P-value from the test of $H_0: var(\beta_c) = 0$	Shrunken estimate of $sd(\beta_c)$	$\frac{\widehat{var(\beta_c)}}{var(\widehat{\beta_c})}$
Worker characteristics:				
Male	-0.003	0.650	0.000	0.0%
Age indicators:				
... 25≤ age ≤ 34	0.051 ***	0.000	0.036	75.3%
... 35 ≤ age ≤ 44	(ref)			
... 45 ≤ age ≤ 54	-0.017 ***	0.477	0.000	0.0%
... 55 ≤ age ≤ 59	-0.031 ***	0.007	0.017	41.9%
Married	-0.008 *	0.153	0.011	20.2%
Native	-0.003	0.824	0.000	0.0%
Number of children:				
... 0	0.029 ***	0.032	0.016	33.8%
... 1	(ref)			
... 2	0.005	0.381	0.004	5.3%
... 3+	0.008 *	0.221	0.009	15.3%
Child below the age of six in household	-0.008 *	0.319	0.007	9.1%
Educational attainment:				
... Primary and lower secondary	-0.003	0.015	0.015	38.4%
... Upper secondary	(ref)			
... Tertiary	0.015 **	0.000	0.023	53.5%
School dropout (any level):	-0.041 ***	0.000	0.025	65.3%
Job Characteristics:				
Index of use of skills at work:				
... ICT	0.002 ***	0.781	0.000	0.0%
... Numeracy	0.000	0.779	0.000	0.0%
... Reading	0.005 ***	0.013	0.003	39.1%
... Writing	0.001	0.028	0.002	34.6%
Managing other employees	0.000	0.236	0.007	14.4%
Occupation (ISCO-based):				
... Skilled	0.007	0.909	0.000	0.0%
... Semi-skilled, white-collar	(ref)			
... Semi-skilled, blue-collar	-0.002	0.693	0.000	0.0%
... Elementary	-0.008 *	0.904	0.000	0.0%
Hourly wage rate, in deciles				
... Lowest decile	0.008	0.704	0.000	0.0%
... 2nd decile	0.004	0.817	0.000	0.0%
... 3rd decile	0.008	0.883	0.000	0.0%

... 4th decile	0.006		0.869	0.000	0.0%
... 5th decile	(ref)				
... 6th decile	0.009		0.508	0.000	0.0%
... 7th decile	0.002		0.804	0.000	0.0%
... 8th decile	-0.005		0.261	0.012	12.9%
... 9th decile	-0.004		0.725	0.000	0.0%
... Highest decile	-0.017	**	0.105	0.019	24.7%
Establishment characteristics:					
Establishment size:					
... 1 to 10 people	0.000		0.034	0.016	33.4%
... 11 to 50 people	-0.004		0.034	0.014	33.3%
... 51 to 250 people	(ref)				
... 251 to 1,000 people	0.000		0.187	0.012	17.7%
... More than 1,000 people	0.010		0.018	0.024	37.2%
Over the last 12 months, the establishment has					
... Increased the number of workers	0.001		0.925	0.000	0.0%
... Stayed more or less the same	(ref)				
... Decreased the number of workers	0.004		0.892	0.000	0.0%
Establishment part of larger firm	0.009	***	0.805	0.000	0.0%
Industry (ISIC):					
... Agriculture, forestry, and fishing (A)	-0.017	**	0.383	0.008	5.2%
... Mfg and other industrial act. (B,C,D,E)	0.000		0.077	0.014	27.3%
... Construction (F)	(ref)				
... Wholesale, retail, food services (G,H,I)	-0.007		0.021	0.018	36.3%
... Information and communication (J)	0.000		0.692	0.000	0.0%
... Financial and insurance activities (K)	0.002		0.008	0.038	41.0%
... Real estate activities (L)	-0.035	***	0.028	0.028	34.6%
... Pro, sci, and admin services (M, N)	0.006		0.015	0.026	38.1%
... Pub adm, edu, and health (O,P,Q)	0.059	***	0.066	0.027	28.6%
... Other service activities (R,S,T,U)	-0.004		0.555	0.000	0.0%

Note: Formal training participation is measured using question *B_ Q04a in* the *PIACC* questionnaire. The country-specific weighted regressions are based on the PIAAC samples of 25–59-year-olds working full-time in the private sector. Model also includes an intercept as well as indicators for item non-response. *** = p<0.01; ** = p < 0.05; * = p < 0.10.

TABLE 7: EFFECTS OF WORKER, JOB, AND ESTABLISHMENT CHARACTERISTICS ON THE PROBABILITY OF INFORMAL TRAINING PARTICIPATION, FULL PIAAC SAMPLE

	Average $\hat{\beta}_c$		P-value from the test of $H_0: var(\beta_c) = 0$	Shrunken estimate of $sd(\beta_c)$	$\frac{\widehat{var(\beta_c)}}{var(\widehat{\beta_c})}$
Worker characteristics:					
Male	0.004		0.001	0.030	48.2%
Age dummies:					
... 25≤ age ≤ 34	0.018	**	0.190	0.017	17.5%
... 35 ≤ age ≤ 44	(ref)				
... 45 ≤ age ≤ 54	-0.015	**	0.663	0.000	0.0%
... 55 ≤ age ≤ 59	-0.050	***	0.321	0.016	9.0%
Married	0.006		0.619	0.000	0.0%
Native	-0.022		0.000	0.055	58.1%
Number of children:					
... 0	0.021	**	0.562	0.000	0.0%
... 1	(ref)				
... 2	0.013		0.295	0.014	10.6%
... 3+	0.006		0.243	0.019	13.9%
Child below the age of six in household	-0.021	**	0.771	0.000	0.0%
Educational attainment:					
... Primary and lower secondary	-0.048	***	0.431	0.007	2.2%
... Upper secondary	(ref)				
... Tertiary	0.038	***	0.007	0.031	42.1%
School dropout (any level):	0.045	***	0.005	0.031	43.1%
Job Characteristics:					
Index of use of skills at work:					
... ICT	0.006	*	0.610	0.000	0.0%
... Numeracy	0.003		0.963	0.000	0.0%
... Reading	0.028	***	0.000	0.009	71.1%
... Writing	0.015	***	0.000	0.007	60.0%
Managing other employees	0.035	***	0.084	0.019	26.5%
Occupation (ISCO-based):					
... Skilled	0.026	***	0.082	0.025	26.7%
... Semi-skilled, white-collar	(ref)				
... Semi-skilled, blue-collar	0.013		0.185	0.022	17.9%
... Elementary	0.000		0.324	0.020	8.8%
Hourly wage rate, in deciles					

... Lowest decile	-0.020		0.271	0.027	12.3%
... 2nd decile	-0.031	**	0.432	0.010	2.2%
... 3rd decile	-0.010		0.138	0.032	21.7%
... 4th decile	-0.005		0.600	0.000	0.0%
... 5th decile	(ref)				
... 6th decile	0.022	*	0.706	0.000	0.0%
... 7th decile	0.027	**	0.924	0.000	0.0%
... 8th decile	0.051	***	0.965	0.000	0.0%
... 9th decile	0.063	***	0.877	0.000	0.0%
... Highest decile	0.077	***	0.835	0.000	0.0%
Establishment characteristics:					
Establishment size:					
... 1 to 10 people	-0.095	***	0.007	0.035	41.6%
... 11 to 50 people	-0.045	***	0.628	0.000	0.0%
... 51 to 250 people	(ref)				
... 251 to 1,000 people	0.034	***	0.546	0.000	0.0%
... More than 1,000 people	0.046	***	0.182	0.025	18.1%
Over the last 12 months, the establishment has					
... Increased the number of workers	0.052	***	0.043	0.023	31.9%
... Stayed more or less the same	(ref)				
... Decreased the number of workers	0.024	***	0.448	0.004	1.2%
Establishment part of larger firm	0.085	***	0.472	0.000	0.0%
Industry (ISIC):					
... Agriculture, forestry, and fishing (A)	0.027		0.001	0.087	48.4%
... Mfg and other industrial act.(B,C,D,E)	0.005		0.000	0.054	52.6%
... Construction (F)	(ref)				
... Wholesale, retail, food services (G,H,I)	0.013		0.004	0.047	43.7%
... Information and communication (J)	0.027		0.676	0.000	0.0%
... Financial and insurance activities (K)	0.087	***	0.079	0.047	27.0%
... Real estate activities (L)	-0.011		0.000	0.139	51.2%
... Pro, sci, and admin services (M, N)	0.034	**	0.152	0.033	20.4%
... Pub adm, edu, and health (O,P,Q)	0.139	***	0.003	0.064	45.5%
... Other service activities (R,S,T,U)	0.068	***	0.015	0.069	38.2%

Note: Informal training participation is measured using question *B_ Q13* in the *PIACC* questionnaire. The country-specific weighted regressions are based on the *PIAAC* samples of 25-59-year-olds working full-time in the private sector. Model also includes an intercept as well as indicators for item non-response. *** = p<0.01; ** = p < 0.05; * = p < 0.10.

Table 8: Published Studies on Impacts of Firm-Based Training

Country/ Study author(s)	Period covered	Data source	Selected sample	Outcome studied	Training measures	Estimator[a]	Impacts
United States							
Lynch (1992)	1980 and 1983	NLSY79	3,064 young workers with wage data in both 1980 and 1983	log hourly wage rate	On-site training: Company-provided training and seminars or training programs at work not run by the employer; weeks completed	OLS, selection, FE	-0.0002
					Off-site programs: business colleges, nursing programs, vocational-technical institutions, barber and beauty schools, flight school, correspondence courses, seminars or training programs outside of work, or vocational rehabilitation; weeks completed.		0.002*
Parent (1999)	1980-1991	NLSY79	5,604 young, private-sector workers	log hourly wage rate	On-site training: Company-provided training and seminars or training programs at work not run by the employer; years completed.	OLS, GLS, IV, IV-GLS	0.12* to 0.17*
					Off-site programs: business colleges, nursing programs, vocational-technical institutions, barber and beauty schools, flight school, correspondence courses, seminars or training programs outside of work, or vocational rehabilitation; years completed while employed.	OLS, GLS, IV, IV-GLS	0.08* to 0.14*

Loewenstein and Spletzer (1999b)	1993 and 1994	NLSY79	4,681 young workers with wage data in 1994	Log hourly wage rate	Training program or on-the-job training designed to improve job skills, help people find a job, or learn a new job, excluding training not partly paid for by employer; number of training spells in current job.	OLS	0.04* to 0.06*
Frazis and Loewenstein (2005)	1979-2000	NLSY79	17,809 job spells with a total of 75,698 observations	log hourly wage rate	Prior to 1987: company-provided training and training spells shorter than one month. After 1987: Training whose explicit cost is at least partly paid for by the employer.	Within-job FE	0.003 (linear); 0.041 (log-version) No t-statistics or s.e. supplied
United Kingdom							
Blundell et al. (1996)	1981-1991	NCDS	2,781 33-year-old employed workers.	wage growth: log hourly wage rate in 1991 minus log hourly wage rate 1981	On-site training.	First-difference OLS	0.04* (males) 0.05* (females)
					Off-site training.		0.07* (males) 0.10* (females)
Booth and Bryan (2005)	1998-2000	BHPS	3,333 16–65-year-old full-time private sector workers	log hourly wage rate	Any skills training, current employer.	FE	0.02*
					Any skills training, previous employer.		0.08*
Dearden et al. (2006)	1985-1996	UK ACP and LFS	94 industry groupings	Log value added per worker and	Any job-connected education or training during past four weeks.	GMM	Elasticity of 0.6% for productivity and 0.3% for

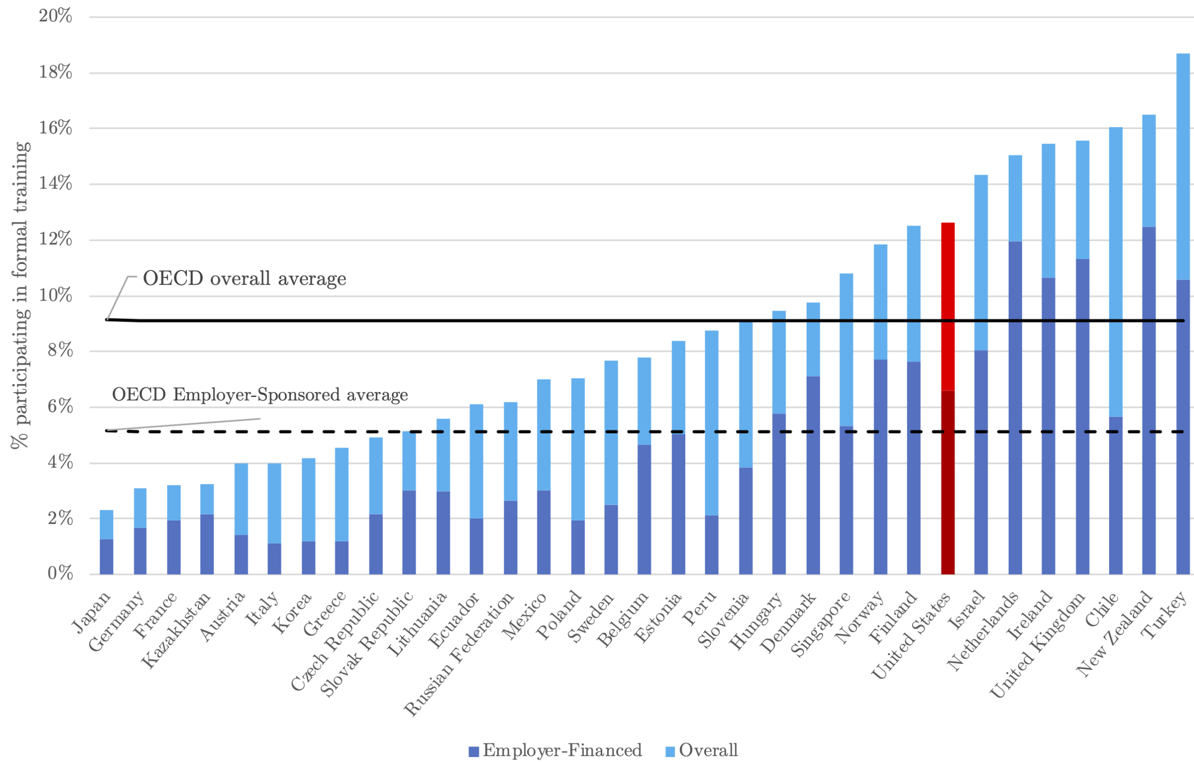
				log hourly wage rate			hourly wage rates
Germany							
Pischke (2001)	1986-1989	GSOEP	3,363 16–64-year-olds working at date of interview, 1986	log monthly earnings	Work-related courses, measured in years.	FE	0.03
Muehler et al. (2007)	2000 and 2004	GSOEP	1,751 full-time employed men with monthly earnings above €600	log monthly wages	Any work-related courses between 2001 and 2003.	PSM	0.05* (ATE) 0.04 (ATT)
Görlitz (2011)	2006-2007	WeLL	5,829 workers employed in one of 149 firms as of Dec. 31, 2006 and still employed by late 2007.	log monthly earnings	Class-room training like courses, seminars or lectures during two-year period prior to interview. Number of attended and planned.	OLS, Tobit	0.05* (no training but intended) to 0.16* (more than three courses)
European Union member states							
Arulampalam et al. (2010)	1994-1999	ECHP	8,826 26-54 year-old private sector workers w/ at least 15 hours of work.	log hourly wages.	Any education or training finished since January of the previous calendar year.	OLS, QR	0.01 (BE, DK) to 0.09* (IR)
Italy							
Brunello et al. (2012)	1999-2005 (bi-annually)	ILFI	1,874 20-55 year old private sector workers.	log monthly earnings.	Weeks of participation in any program organized by firms, local authorities, or industrial organization during the previous year.	OLS, RE, IV, RE-IV	0.6% (OLS) to 1.4% (RE-IV)
Finland							
Jones et al. (2012)	2000-2004	Administrative data	223 local co-operative banks	Average annual wage.	Training expenditures and training duration separated	FE	0.07% (expenditures)

					into training offered centrally and training purchased locally.		and 0.03% (duration)
The Netherlands							
De Grip and Sauermann (2012)	32 weeks in 2008-2009	Administrative data	74 call center workers .	Average time to handle a call.	One-week course on techniques to decrease time to handle calls without lowering call quality.	RCT, OLS, FE	0.01 (unconditional) to 0.12 (conditional)

NOTE: [a] Characterizes the class of estimator used in the analysis. "OLS" refers to a standard linear regression. "Selection" refers to a parametric econometric model of selection bias that controls for the correlation between unobservables in the outcome and participation equation (Heckman 1976, 1979). "FE" denotes a fixed effects estimator. "RE" denotes a random effects model. "IV" denotes an instrumental variables regression. "GMM" denotes a generalized methods of moment estimator. "QR" refers to a quantile regression. "PSM" denotes a propensity score matching estimator. "Tobit" refers to a maximum likelihood method used on censored data. "RCT" refers to a randomized control trial. * indicates a statistically significant estimates on at least a 5% level.

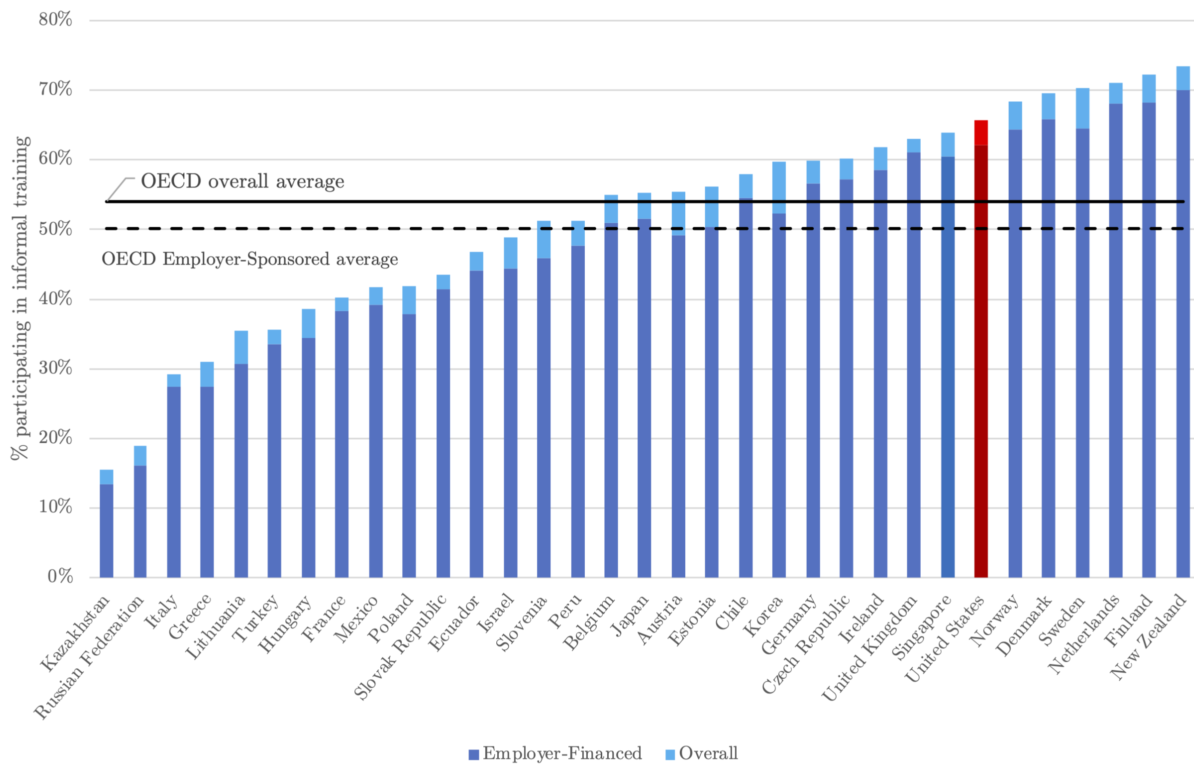
10 Figures

Figure 1: Formal Training Incidence Across Countries



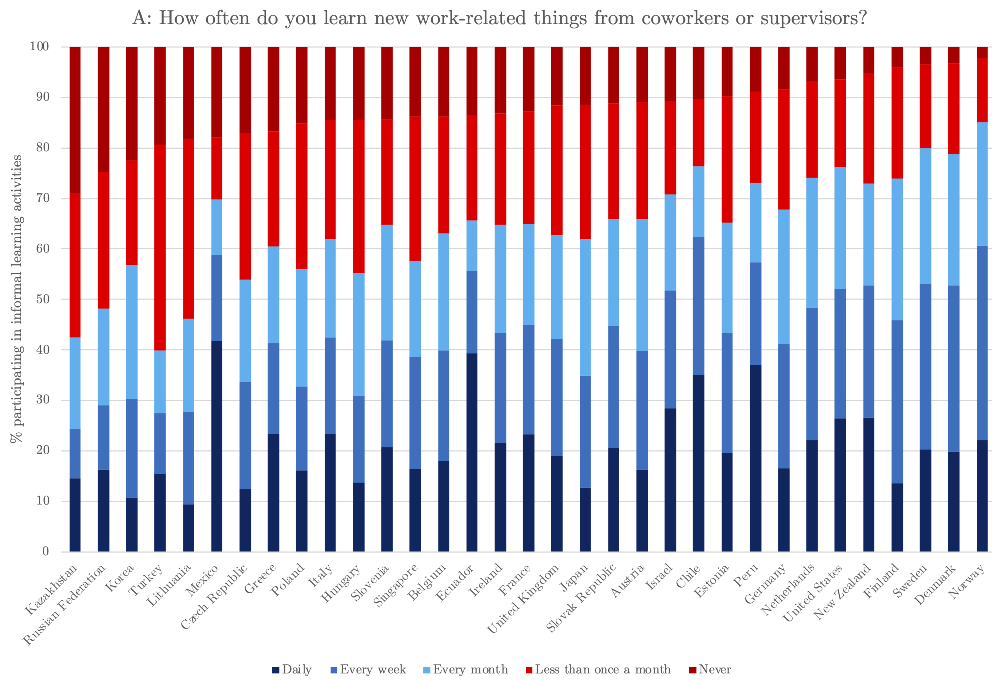
Note: The formal training measure is taken from the *PIAAC* sample of 25-59-year-olds working full-time in the private sector. OECD averages are (unweighted) country averages of available OECD member states.

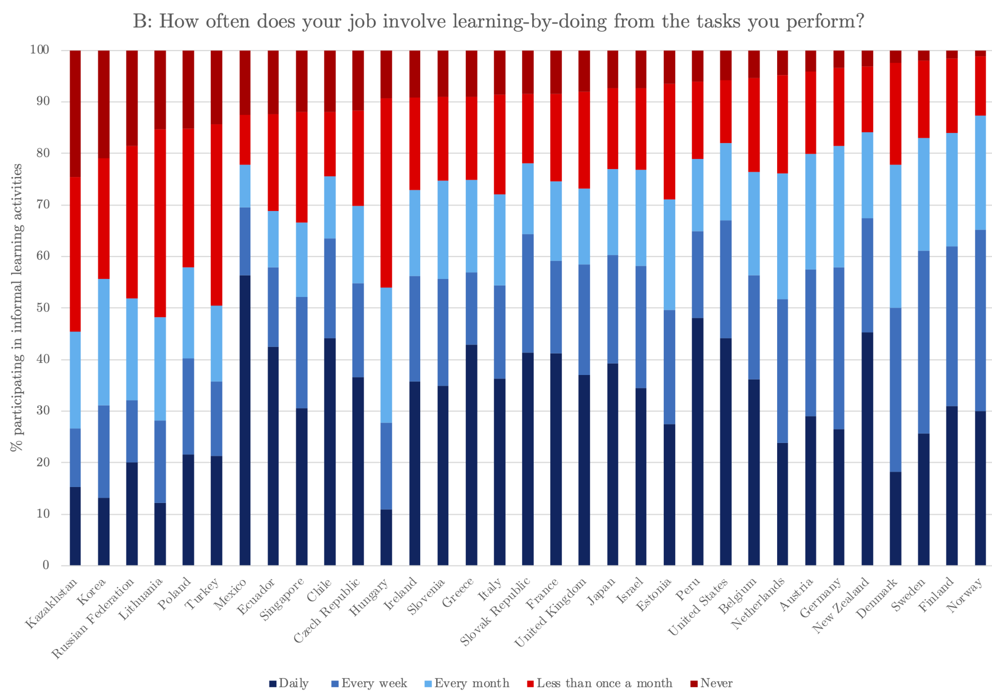
Figure 2: Informal Training Incidence Across Countries



Note: The informal training measure is taken from the *PIAAC* sample of 25-59-year-olds working full-time in the private sector. OECD averages are (unweighted) country averages of available OECD member states.

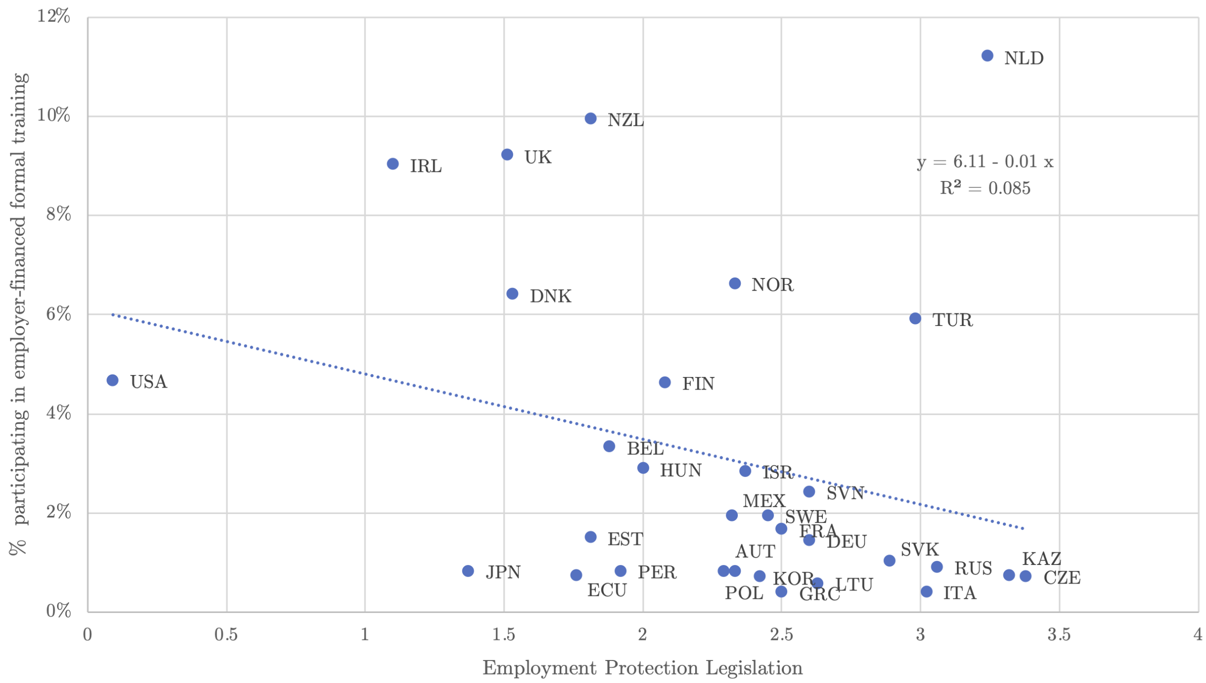
Figure 3: Casual training and learning-by-doing across countries





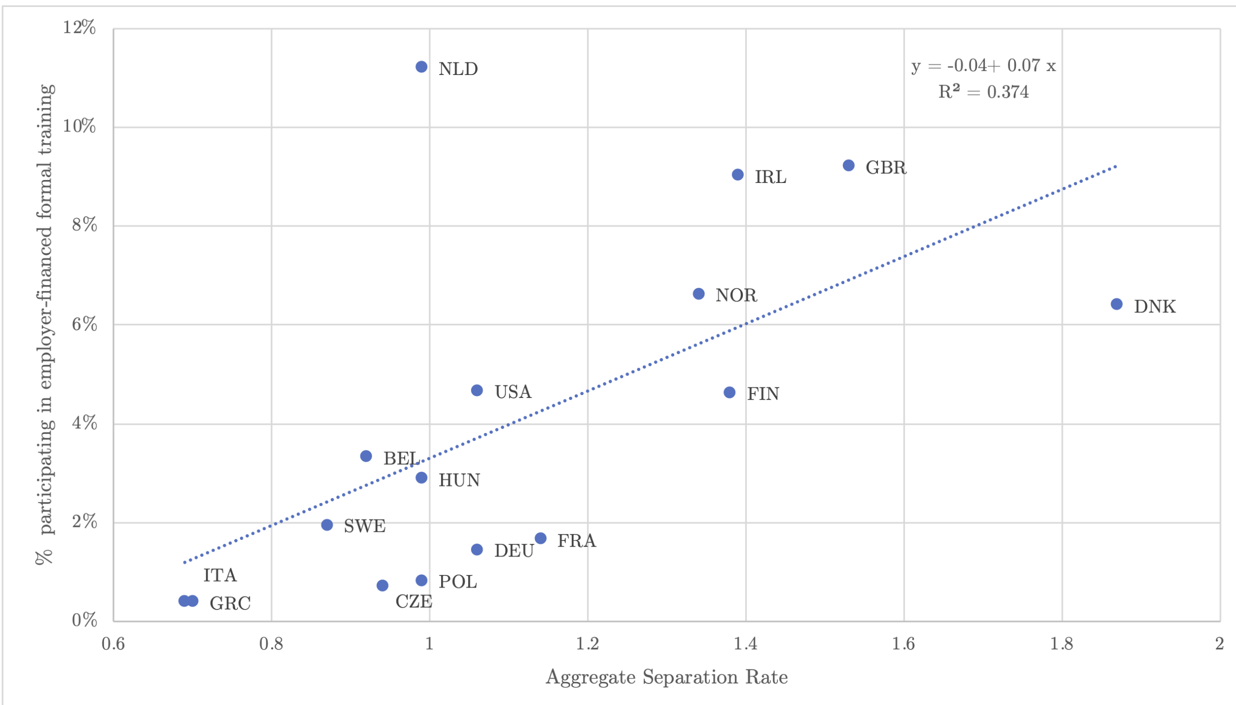
Note: The casual training and learning-by-doing measures are taken from the *PIAAC* sample of 25-59-year-olds working full-time in the private sector. OECD averages are (unweighted) country averages of available OECD member states.

Figure 4: Employment Regulation and Employer-Financed Formal Training Across Countries



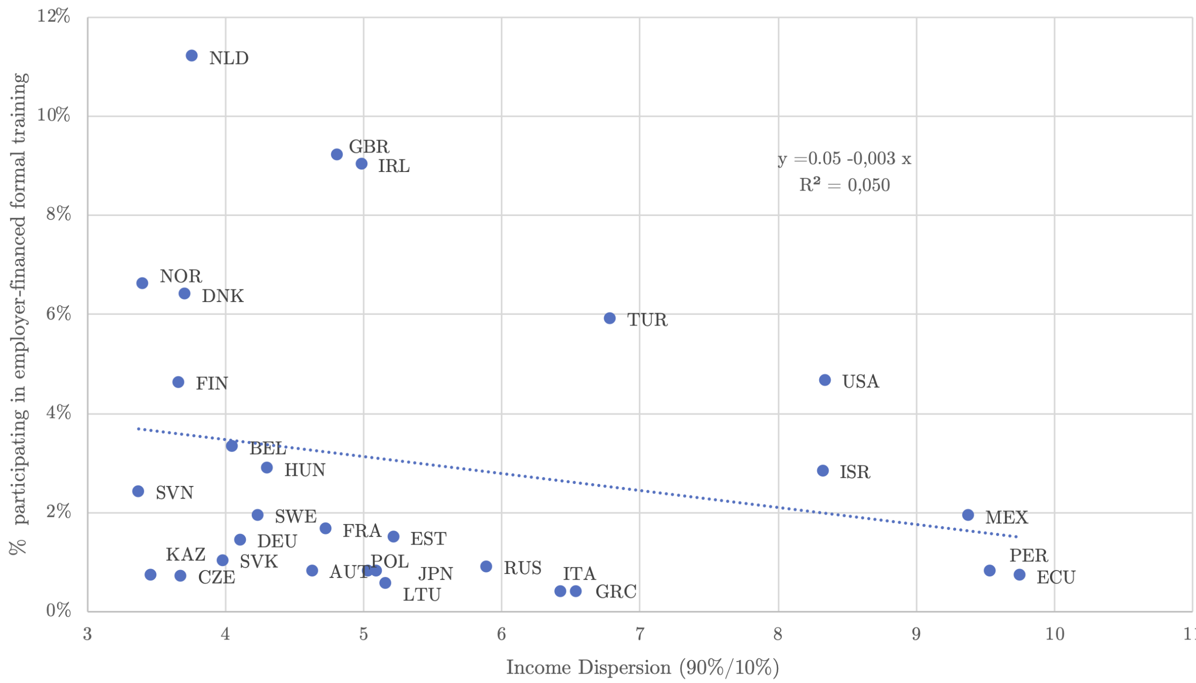
*Note:*Employment Protection Legislation (EPL) is the OECD's measure of strictness of employment protection - individual dismissals (regular contracts). The underlying scale runs from zero to six with higher scores representing stricter regulation (https://stats.oecd.org/Index.aspx?DataSetCode=EPL_R). EPL is measured in 2011. The participation measure comes from the answer to PIAAC question B_Q11 and is based on full-time, private-sector workers ages 25 to 59. Formal training is measured in 2011. The empirical correlation equals -0.29.

Figure 5: Separation Rates and Employer-Financed Formal Training Across Countries



Note: The monthly aggregate separation rate comes from Table 2 in Hobijn and Şahin (2009). Separation rates are measured during 1990 to 2006. The participation measure comes from the answer to PIAAC question B_Q11 and is based on full-time, private-sector workers ages 25 to 59. Formal training is measured in 2011. The empirical correlation equals 0.61.

Figure 6: Income Compression and Employer-Financed Formal Training Across Countries



Note: The income dispersion measure comes from Roser and Ortiz-Ospina (2013). Income dispersion is measured in 2011 except for Japan (2008), the US (2010), Israel (2012), and Mexico (2012). Consumption rather than income is used for Kazakhstan, Russia, and Turkey. The participation measure comes from the answer to PIAAC question B_Q11 and is based on full-time, private-sector workers ages 25 to 59. Formal training is measured in 2011. The empirical correlation equals -0.22.