CESifo WORKING PAPERS

8288
2020

May 2020

Unemployment and Endogenous Reallocation over the Business Cycle

Carlos Carrillo-Tudela, Ludo Visschers

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. 8288*

Unemployment and Endogenous Reallocation over the Business Cycle

Abstract

This paper studies the extent to which the cyclicality of gross and net occupational mobility shapes that of aggregate unemployment and its duration distribution. Using the SIPP, we document the relation between workers' (gross and net) occupational mobility and unemployment duration over the long run and business cycle. To interpret this evidence, we develop an analytically and computationally tractable stochastic equilibrium model with heterogenous agents and occupations as well as aggregate uncertainty. The model is quantitatively consistent with several important features of the US labor market: procyclical gross and countercyclical net occupational mobility, the large volatility of unemployment and the cyclical properties of the unemployment duration distribution, among others. Our analysis shows that "excess" occupational mobility due to workers' changing career prospects interacts with aggregate conditions to drive fluctuations of aggregate unemployment and its duration distribution.

JEL-Codes: E240, E300, J620, J630, J640.

Keywords: unemployment, business cycle, rest, search, occupational mobility.

Carlos Carrillo-Tudela
Department of Economics
University of Essex
United Kingdom – Colchester, CO4 3SQ
cocarr@essex.ac.uk

Ludo Visschers
School of Economics
The University of Edinburgh
United Kingdom – Edinburgh, EH8 9JT
Ludo.Visschers@ed.ac.uk

April 2020
This paper owes much to many. We would like to thank the editor and four referees for their insightful comments. We have also benefited from comments and suggestions of many colleagues in workshops and seminars. We would like to thank participants of the CESifo workshop "Labor Market Search and Policy Applications", the Labor Market Dynamics and Growth conferences, the NBER Summer Institute (RSW), VI REDg - Barcelona, SED Meetings and SaM workshops, the CESifo Macroeconomics and Survey Data conference, the Kiel/Richmond Fed Workshop, and in seminars at Oxford, Carlos III, Essex, the St. Louis Fed, CEMFI (MadMac), Konstanz, Surrey, VU Amsterdam, UCL, Queen Mary, Edinburgh, Tinbergen Institute, Sheffield, Uppsala, Tel Aviv, Toronto, Western Ontario, Southampton, UC Louvain, Cergy-Pontoise, Frankfurt, Bonn, Cologne, Tilburg, Oslo, BI Oslo, Bristol, Cerge-EI, Einaudi, Autonoma de Barcelona, Sveriges Riksbank, IIES Stockholm, European University Institute and Nottingham. We thank Rahel Felder and Junko Iwagami who provided excellent research assistance with the data analysis. The usual disclaimer applies. Carlos Carrillo-Tudela acknowledges financial support from the UK Economic and Social Research Council (ESRC), award reference ES/I037628/1. Ludo Visschers acknowledges financial support from the Juan de la Cierva Grant and Project Grant ECO2010-20614 from the Spanish Ministry of Science and Innovation, a grant from the Bank of Spain's Programa de Investigación de Excelencia, and the MacCaLM project at the University of Edinburgh (ESRC award reference ES/L009633/1). He also thanks the CESifo institute and the St. Louis Fed for their hospitality. First version: June 2010.

1 Introduction

This paper studies the role of gross and net occupational mobility in shaping aggregate unemployment and its duration distribution over the business cycle. In the data occupational mobility and unemployment are closely intertwined. We document that on average 44% of workers who went through a spell of unemployment in the US changed "major occupational groups" at re-employment.[1] These occupational movers also take 18% longer to find a job than occupational stayers. Crucially, in recessions unemployment durations of occupational movers increase more than those of stayers, contributing to the countercyclical increase in long-term unemployment. Recessions are also times when occupations which in net gain workers through unemployment do so more. Thus it appears that recessions are times in which occupation-wide differences play a larger role in reallocation and unemployment. However, accompanying the countercyclical increase in net mobility and lengthening of unemployment spells, there is a decrease in the overall proportion of unemployed workers switching occupations, suggesting that recessions are worse times to move to a better occupation.

To investigate the economic mechanism that links the above gross and net mobility patterns to cyclical unemployment outcomes, we propose and quantitatively assess a tractable equilibrium business cycle model in which the unemployed face search frictions in, and reallocation frictions across, heterogeneous labour markets. We connect with two seminal approaches in the literature. In the Lucas and Prescott (1974) "islands" framework countercyclical net mobility across sectors (or "islands") is typically the main underlying force behind unemployment fluctuations (see Lilien, 1982, Rogerson, 1987, Gouge and King, 1997). Countercyclical unemployment can arise when more workers engage in time consuming switches from sectors that have been affected harder in a recession to those sectors which offer relatively higher job finding prospects. An alternative approach follows the canonical search and matching model, as described by Mortensen and Pissarides (1994), which explains unemployment fluctuations through movements in the aggregate job finding and separation rates. In recessions more employed workers lose their jobs and face lower job finding prospects as more matches become unproductive and firms create less vacancies. This framework abstracts from worker mobility across occupations or industries, but summarises any frictions that impede the pairing of jobs and workers across the economy through an aggregate matching function.

Our theory and quantitative analysis builds on these two approaches, but differently, we argue that the interaction of stochastic aggregate productivity and worker's own career prospects within an occupation (which drive "excess" occupational mobility) is the main channel behind the fluctuations of aggregate unemployment and its duration distribution.[2] The key mechanism at play is that with uncertain returns and costly reallocation, unemployed workers are more likely to wait for their labor

[1] Major occupational groups are broad categories that can be thought of as representing one-digit occupations. For example, managers, sales, mechanic and repairers, construction/extraction, office/admin support, elementary trades, etc. The above proportion is obtained after correcting for measurement error.

[2] This result echoes the conclusion of Jovanovic and Moffit (1990), who show that matching considerations (excess mobility) are more important than sectoral shocks in explaining labor mobility through direct job-to-job transitions in an aggregate steady state environment. In the data, excess occupational mobility arises when workers move occupations in ways that offset each other to yield constant occupational sizes, and represent about 90% of gross mobility among the unemployed.

market prospects to improve in their pre-separation occupations rather than search for jobs in new occupations during downturns, even though some occupations offer on average relative better employment prospects than others. In expansions the opposite happens, rendering gross occupational mobility procyclical. This margin is quantitatively very important over the business cycle and creates a strong amplification force in cyclical unemployment and in a wide range of other aggregate labor market variables. During recessions a smaller proportion of unemployed workers whose career prospects have deteriorated sufficiently to find it profitable to change occupations, weigh more heavily differences in occupation-wide productivities to create countercyclical net mobility.

Although net mobility has a small role in explaining unemployment fluctuations, it has a clear cyclical pattern. During recessions a higher proportion of workers lose their jobs in routine manual occupations and do not come back to these jobs; while a higher proportion of workers find jobs in non-routine manual occupations at re-employment. We show that these patterns contribute to the long-run decline of the employment share of routine occupations and long-run increase in the employment share of non-routine occupations (see also Cortes et al., 2020). We find no contradiction between changing career prospects playing a very important role in shaping cyclical unemployment, and worker flows through unemployment contributing meaningfully to the changing sizes of occupations particularly during recessions.

A large literature has developed further the ideas of Lucas and Prescott (1974) and Mortensen and Pissarides (1994) to investigate unemployment fluctuations.[3] This literature, however, typically finds it elusive to quantitatively reproduce the observed cyclical patterns of gross and net mobility and/or the observed cyclical patterns of aggregate unemployment. A key contribution of our framework is that it is able to quantitatively reproduce not only a wide range of time-aggregated occupational mobility and unemployment patterns, but simultaneously generates the empirical procyclicality of gross occupational mobility and the countercyclicality of net occupational mobility among the unemployed along with the cyclical fluctuations of aggregate unemployment and its duration distribution. Underlying these fluctuations, the cyclical responses of the model's aggregate job separation and job finding rates are in line with the data. The model also generates a high correlation between the job finding rate and labor market tightness and a strongly downward-sloping Beveridge curve.

The economy we consider exhibits both aggregate and occupational-wide productivity fluctuations, where the latter drives net occupational mobility. In addition, idiosyncratic worker-occupation productivity shocks, orthogonal to the occupation-wide shocks, capture the evolving career prospects of a worker within an occupation and drive excess mobility. Workers also accumulate occupation-specific human capital through learning-by-doing, but face skill loss during an unemployment spell. Even with this rich level of heterogeneity, workers' job separations and occupational mobility decisions can be characterised by simple reservation productivity cutoffs that respond to aggregate and occupational-wide conditions.

These cutoffs imply that a jobless spell has different phases. If the job separation cutoff is above

[3] In addition to the aforementioned studies see also Abraham and Katz (1986), Hamilton (1988), Shimer (2005), Hall and Milgrom (2008), Veracierto (2008), Hagedorn and Manovskii (2008), Dvorkin (2014) and Pilossoph (2016), among many others.

the reallocation cutoff, workers can transit between three phases. A *search unemployment* phase occurs when a worker's occupational productivity is above the separation cutoff as firms post jobs, but frictions make it time-costly to find a vacancy. When this worker's productivity deteriorates and lies between the separation and reallocation cutoffs, he transits to a *rest (or wait) unemployment* phase, as concurrently no employment opportunities exist for him in his pre-separation occupation. As his job prospects deteriorates further such that his productivity falls below the reallocation cutoff, the worker enters a *reallocation unemployment* phase, but only during the time it takes to find a different occupation in which his productivity is above the reallocation cutoff. On the other hand, if the separation cutoff lies below the reallocation cutoff, workers can transit only between two phases: search and reallocation unemployment. As these two cases deliver very different empirical implications, we use the observed behavior of occupational mobility through unemployment and unemployment duration to distinguish between them.

The study of occupational (or industry) mobility focused exclusively on workers who went through unemployment, however, has received relatively little attention compared to the larger amount of research investigating occupational mobility among pooled samples of employer movers and stayers (see Jovanovic and Moffitt, 1990, Kambourov and Manovskii, 2008, and Moscarini and Thomsson, 2007, among others).[4] The same applies even more strongly to the cyclicality of unemployed workers' occupational mobility and its relation with unemployment duration. There is no reason, a priori, to conclude that their gross and net mobility patterns should have the same cyclical properties as the ones already documented for pooled samples. We use data from the Survey of Income and Programme Participation (SIPP) between 1983-2014 to derive the relationship between occupational mobility and unemployment duration. We use this relationship to investigate the degree of attachment unemployed workers have to their pre-separation occupations and how it evolves with their spell duration. We also investigate how this attachment differs across demographic groups, occupational categories and across unemployment spells, how it depends on excess and net mobility and the business cycle. We further use the Panel Survey for Income Dynamics (PSID) to corroborate some of our results (see also Carrillo-Tudela et al., 2014, for an analysis using the Current Population Survey (CPS)).

When constructing these occupational mobility statistics an important concern is the extent to which coding errors creates spurious mobility. We show that occupational miscoding increases observed excess mobility and reduces the importance of net mobility (see also Kambourov and Manovskii, 2013). It also makes occupational mobility appear less responsive to unemployment duration and the business cycle. Crucially, we show that one cannot use existing correction estimates based on samples pooling all workers when attempting to correct the occupational mobility of the unemployed. As these are key dimensions of our analysis we develop a novel classification error model that allows us to estimate the extent of coding error at the level of each occupation. We take advantage of a change in survey design in the SIPP to estimate coding errors and use these estimates to clean our data.

We calibrate our model using simulation method of moments. The calibration implies that search,

[4] A few recent exceptions are Şahin et al. (2014), Fujita and Moscarini (2017), Carrillo-Tudela et al. (2016), Wiczer (2015), Huckfeldt (2016), Belot et al. (2018) and Faberman and Kudlyack (2019).

rest and reallocation unemployment episodes can arise during a jobless spell. A key insight is that recessions are characterised by a wider difference between the separation and reallocation cutoffs relative to expansions and this occurs in all occupations, irrespectively of cyclical changes in occupational-wide productivity differences. This fanning out of the cutoffs also implies that the nature of unemployment changes over the cycle: during recessions workers spend a longer period of their jobless spells as rest unemployed waiting for conditions to improve in their occupation. As aggregate conditions improve workers leave rest unemployment faster by crossing either the separation or the reallocation cutoff and begin spending most of their jobless spell in search unemployment episodes. These dynamics result in procyclical excess (and gross) occupational mobility and are behind the observed fluctuations of aggregate unemployment and other aforementioned labor market variables. Interacted with occupational-wide productivity differences they also play an important role in shaping the cyclical responses of unemployment in each occupation as well as these occupations' outflows and inflows, such that the model is able to replicate each occupations' observed cyclical net mobility patterns.

Related Literature

A key component of our analysis is that workers are heterogeneous in their occupation-specific productivities. A growing literature incorporates time-invariant worker heterogeneity to the Diamond-Mortensen-Pissarides (DMP) framework to investigate its cyclical properties (see Bils et al., 2012, Chassambouli, 2013 and Murtin and Robin, 2018, among others). As we are interested in the roles of gross and net mobility in determining unemployment fluctuations, worker heterogeneity in our model is naturally time variant. Nevertheless, part of the mechanism through which our model generates amplification in cyclical unemployment is shared with some of these papers. Namely, some unemployed workers do not provide incentive for vacancies to enter the labor market during periods in which their productivities lie below the separation cutoff. Allowing for occupational mobility, however, remains crucial to generate the observed fluctuations of aggregate unemployment and its duration distribution. In this context, the combination of aggregate shocks and time-variant worker heterogeneity also allows the model to generate an increased difference between the unemployment durations of occupational movers and stayers during recessions.

Re-estimating a version of our model without occupational mobility shows that its cyclical performance is drastically reduced in many important dimensions. We show that although it is possible to improve in some of these dimensions, for example the volatility of aggregate unemployment, one needs to create too much long-term unemployment in normal times. Hence, one has to be willing to give up on fitting the long-run and cyclical moments that characterise the unemployment duration distribution. Even with this caveat, the occupational mobility model remains able to consistently provide a much better fit to the remainder empirical long-run and cyclical moments.

Alvarez and Shimer (2011) use evolving industry productivities to study the importance of rest unemployment in determining the level of aggregate unemployment. They consider a steady state economy with different industries each of which is characterised by a competitive labor market. In this environment workers are indifferent between work and rest or between work, rest and reallocation. This implies that the overall size of employment, rest and reallocation unemployment is determined, but not the individual transitions between these categories. In contrast, in our model the stochastic process that drives workers' career prospects within an occupation determines the transitions of an individual worker between employment, search, rest and reallocation unemployment. This difference implies that our model is suitable to analyse the relationship between unemployment duration, occupational mobility and job finding probabilities, both in the long-run and over the cycle.

Kambourov and Manovskii (2009a) investigate the rise of wage inequality in the US using a steady state model in which workers accumulate occupation-specific human capital and search randomly across occupations. In their model, occupation idiosyncratic productivity shocks and human capital accumulation also yield increased attachment to an occupation, where the latter introduces a moving cost as it is lost when changing occupations. There, as in Alvarez and Shimer (2012) and Rogerson (2005), these features generate a waiting motive that is similar to the one described here. Our paper, however, focuses on the unemployed, explicitly considers both net and gross mobility and studies the business cycle. In particular, differences in human capital across workers help our model reproduce the cyclical differences in the unemployment patterns between young and prime-aged workers. In particular, the proportion of young and prime-aged workers in aggregate unemployment and job separations changes cyclically. During recessions the composition of unemployment and separations moves towards the (on-average) more productive group of prime-aged workers (see Mueller, 2017).

Jovanovic (1987) introduced rest/wait unemployment in a cyclical setting, but did not link it to occupational mobility, the unemployment duration distribution, or investigated its quantitative properties. More recently, Wiczer (2015) also analysed the role of occupation-wide shocks on long-term unemployment over the cycle. In contrast, unemployed workers in our model take into account the potential recovery of their occupational productivities when making job separations and occupational mobility decision. This gives rise to episodes of rest unemployment and takes us a long way in replicating the overall volatility of cyclical unemployment, while remaining consistent with the cyclical behavior of the long-term unemployed.

The rest of the paper proceeds as follows. Section 2 presents the empirical evidence that motivates our paper. Section 3 describes and characterises the model and its main implications. In Sections 4 and 5 we quantitatively assess this model and show the importance of changing career prospect in explaining cyclical unemployment outcomes. Section 6 concludes. All proofs, detailed data, quantitative analysis and robustness exercises are relegated to several appendices and an online Supplementary Appendix.

2 Occupational Mobility of the Unemployed

Our main statistical analysis is based on the sequence of 1984-2008 SIPP panels, covering the 1983-2014 period. The sample restricts attention to those workers who were observed transiting from employment to unemployment and back within a given panel, and excludes those in self-employment, in the armed forces, or in the agricultural occupations. In our baseline analysis we consider workers who have been unemployed throughout their non-employment spells, but show that our main results also hold when using mixed unemployment/out-of-labor-force spells. An individual is considered unemployed if he/she has not been working for at least a month after leaving employment and reported "no job - looking for work or on layoff". Since we want to focus on workers who have become unattached from their previous employers, we consider those who report to be "with a job - on layoff", as employed.[5]

After dropping all observations with imputed occupations, we compare each workers' reported occupations before and after the non-employment spell. To capture meaningful career changes we use the 21 "major" occupational groups of the 2000 Census Occupational Classification (2000 SOC) as well as their aggregation into the task-based occupational categories proposed by Autor and Dorn (2013) and Cortes et al. (2020). In the SIPP, however, the occupation information of a worker newly hired from unemployment is collected under independent interviewing, which is known to generate occupational coding errors.[6] Miscoding an occupation just before or after an unemployment spell can thus turn an occupational stayer into an observed occupational mover. Without correcting for miscoding we could potentially be inflating the importance of occupational mobility among the unemployed. We address this issue using several approaches to show that our conclusions are very robust.

In the next subsection we develop the classification error model, our main approach to clean the data. The adjustments based on this model are in line with our alternative approaches involving retrospective self-reported mobility in SIPP topical modules, retrospectively-recoded occupations in the PSID and a mobility measure based on the simultaneous movement across industry and occupation. We then present the long-run and business cycle facts that motivate our paper. In Supplementary Appendices A to F we present a more detailed analysis of our long-run and business cycle findings, as well as extensive robustness exercises using the SIPP and PSID and a full analysis of our alternative approaches to correct for classification error across occupations. In Supplementary Appendix G we provide details on the data construction and measurement.

2.1 Correcting for Coding Errors in Occupation Mobility

A number of studies propose different methodologies to correct for coding errors. Some estimate the relative adjustment needed to correct the average occupational mobility rate of a pooled sample of employer movers and stayers (see Kambourov and Manovskii, 2008). However, these estimates

[5]Fujita and Moscarini (2017) find that the unemployed (as typically defined by the BLS) consist of two groups that behave very differently: "temporary laid-off workers" and "permanent separators". The latter group are those who lost their job with no indication of recall. Similarly, Hornstein (2013) and Ahn and Hamilton (2018) consider two groups among the unemployed in terms of fixed characteristics: those with "high job finding rates" and those with "low job finding rates". Excluding from our unemployment measure those workers who are "with a job - on layoff" and those who find employment within a month means that our unemployment sample is close to Fujita and Moscarini's "permanent separators" sample and to Hornstein's and Ahn and Hamilton's "low job finding rate" workers. In Supplementary Appendix D.4.1, we further discuss this issue.

[6]This implies that the occupational question is asked without reference to the answers giving by the respondent in previous interviews. A professional coder then assigns an occupational code based on the respondent's answer, also without reference to occupational codes previously assigned or previous answers given by the same respondent.

do not automatically carry over to the unemployed. We show that coding errors have a much larger absolute and relative impact on the occupational mobility rates of employer stayers (the majority of workers) than on the mobility rates of employer movers.[7]

Instead we propose and implement a classification error model that allows us to identify and estimate the probabilities that an occupation i gets miscoded as another occupation j. This approach has four main advantages. It captures both coder and respondent error. It is able to identify and correct for the large heterogeneity in (the propensity of) coding errors found in the flows between particular occupations. As it relates the observed occupational codes to the underlying true occupations, it allows for a cleaner analysis of net mobility, where the true identity of occupations matter, and of repeat mobility. Finally, it can be easily incorporated in our quantitative analysis. Appendix A describes in detail this correction methodology. Here we briefly discuss its setup and show the main results.

Suppose that coding errors are made according to a garbling matrix $\mathbf{\Gamma}$ of size OxO, where O denotes the number of occupational categories. The element γ_{ij} is the probability that the true occupation $i = 1, 2, ..., O$ is coded as occupation $j = 1, 2, ..., O$, such that $\sum_{j=1}^{O} \gamma_{ij} = 1$. Let $\mathbf{M}$ denote the matrix that contains workers' *true* occupational flows, where element m_{ij} is the flow of workers from occupation i to occupation j. Under independent interviewing such a matrix appears as $\mathbf{M^I} = \mathbf{\Gamma}'\mathbf{M}\mathbf{\Gamma}$, where the pre- and post-multiplication by $\mathbf{\Gamma}$ takes into account that the observed occupations of origin and destination would be subject to coding error. Knowledge of $\mathbf{\Gamma}$ (and of its invertibility) allows us to de-garble $\mathbf{M}$ as $\mathbf{\Gamma^{-1}}'\mathbf{M^I}\mathbf{\Gamma^{-1}}$.

This formulation builds on Poterba and Summers (1986) and Abowd and Zellner (1985), who focus on miscoding of labor force status. They are able to directly observe miscoding from CPS re-interviews, where discrepancies in labor force status are explicitly reconciled by the Census, under the assumption that re-interviews uncovers the true worker's status. In contrast, our challenge is that we do not observe the garbling matrix of occupations directly from the data. Proposition 1 (below), however, shows that it is possible to estimate $\mathbf{\Gamma}$ from the SIPP by using a change in survey design.

To identify and estimate $\mathbf{\Gamma}$ we make three assumptions. (*A1*) *Independent classification errors*: conditional on the true occupation, the realization of an occupational code does not depend on workers' labor market histories, demographic characteristics or time. (*A2*) *"Detailed balance" in miscoding*: coding mistakes are symmetric in that the number of workers whose true occupation i gets mistakenly coded as j is the same as the number of workers whose true occupation j gets mistakenly coded as i. (*A3*) *Strict diagonal dominance*: It is more likely to correctly code occupation i than to miscoded it. The first assumption is also present in Poterba and Summers (1986) and Abowd and Zellner (1985) and is consistent with independent interviewing.[8] The second assumption implies that the overall size of occupations do not change with coding error (see Keane and Wolpin, 2001). The

[7] We also cannot use the common approach that classifies occupational transitions as genuine when the observed occupational change is accompanied by an employer or position change (see Neal, 1999). Following this approach would effectively assume a minimum impact of coding error on the occupational mobility of the unemployed.

[8] A standard practise when using independent interviewing is that professional coders are only given the respondents' contemporaneous job descriptions to assign occupational codes. Mathiowetz (1992) shows that coder error is the main source of classification error when coding occupations. She shows that the importance of coder error is two to five times larger than the importance of the respondent error, depending on the level of occupational code aggregation.

third assumption implies $\gamma_{ii} > 0.5$ (see Hausman et al., 1998) and is verified in our data.

To implement our correction method we exploit the change from independent to dependent interviewing that occurred between the 1985 and 1986 SIPP panels. From the 1986 panel onwards, employed workers were asked whether their work activities or employers had changed relative to the previous interview. Respondents who reported a change in employer or a change in their main activities without an employer change were asked "independently" to describe their occupation. If respondents declared no change in employer *and* in their main activities, the occupational code assigned to the respondent in the previous interview was carried forward. Since the 1985 and 1986 panels overlap between February 1986 and April 1987, we use data from this period to measure occupational flows for the same underlying population under the two survey designs.

Suppose that we were to subject this underlying population to dependent interviewing as applied in the 1986 panel. The occupational flow matrix of the population would then be given by $\mathbf{M^D} = \mathbf{\Gamma}'\mathbf{M_s} + \mathbf{\Gamma}'\mathbf{M_m}\mathbf{\Gamma}$, where $\mathbf{M_s}$ is the matrix that contains the *true* occupational flows of employer/activity stayers and $\mathbf{M_m}$ the matrix that contains the *true* occupational flows of employer/activity changers. Note that only the employer/activity changers are coded independently under this design and hence $\mathbf{M^I_m} = \mathbf{\Gamma}'\mathbf{M_m}\mathbf{\Gamma}$. If instead we were to subject this same population to independent interviewing as applied in the 1985 panel, their occupational flow matrix would be given by $\mathbf{M^I} = \mathbf{\Gamma}'\mathbf{M_s}\mathbf{\Gamma} + \mathbf{M^I_m}$. In Appendix A we show that $\mathbf{\Gamma}$ is identified from $\mathbf{M^I_s} = \mathbf{M^I} - \mathbf{M^I_m}$ and that it can be estimated using the information contain in the 1985 and 1986 panels.

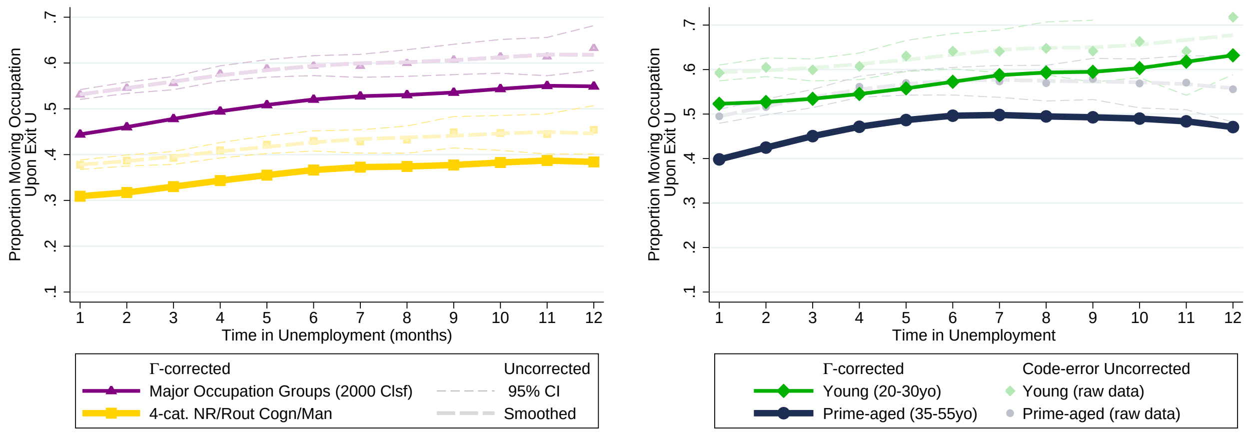
Proposition 1. *Given (A1)-(A3),* $\mathbf{M^I_s} = \mathbf{M^I} - \mathbf{M^I_m}$ *identifies* $\mathbf{\Gamma}$. *Further,* $\mathbf{\Gamma}$ *can be consistently estimated from* $\hat{\mathbf{M}}^{\mathbf{I,85}}$ *and* $\hat{\mathbf{M}}^{\mathbf{I,86}}_{\mathbf{m}}$*, where* $\hat{\mathbf{M}}^{\mathbf{I,85}}$ *and* $\hat{\mathbf{M}}^{\mathbf{I,86}}$ *are sample estimates of* $\mathbf{M^I}$ *based on the 1985 and 1986 SIPP panels.*

It is crucial for this identification approach that we observe, among all employer/activity changers, those who change occupational code but also those who stayed with the same code. A simple mobility rate comparison between the 1985 and 1986 panel would not suffice to identify $\mathbf{\Gamma}$ or even average miscoding, as we need to know the extent of independent coding in both panels. Applying the $\mathbf{\Gamma}$-correction to the occupational flows of workers who go through unemployment and using the "major" occupational categories of the 2000 SOC, results in an average miscoding of about 10% each time information is collected. This implies that at re-employment true occupational stayers have on average about a 20% chance of appearing as occupational movers. Further, we find that different occupations have very different propensities to be assigned a wrong code and, given a true occupation, some coding mistakes are much more likely than others. Statistics that involve occupational identities or directions of occupational flows, as a result, will be affected heterogeneously by miscoding.

We also find that the $\mathbf{\Gamma}$-correction captures very well coding error in the data and also yields good out-of-sample predictions. Moreover, the $\mathbf{\Gamma}$-correction implies an average occupational mobility rate at re-employment that is in line with the one derived from the PSID retrospective occupation-industry supplementary data files. As shown below it is also consistent with two alternative measures of occupational mobility: (i) simultaneously mobility of major occupational and major industrial groups

Figure 1: Extent of occupational mobility by unemployment duration

(a) Overall (b) Age groups - Major occupations



at re-employment and (ii) self-reported duration of occupational tenure obtained from the topical modules of the SIPP.

2.2 Gross Occupational Mobility and Unemployment Duration

We now investigate the extent of occupational mobility of the unemployed and how it changes with unemployment duration. This will allow us to gauge the degree of "attachment" workers have to their pre-separation occupation as their unemployment duration increases. For brevity we will use the terms 'occupational mobility' and 'occupational mover/stayer' to refer exclusively to those workers who go through unemployment. In Figure 1 we pool the SIPP panels to generate aggregate occupational mobility – unemployment duration profiles (henceforth mobility-duration profiles). They show, for a given unemployment duration x, the proportion of workers who changed occupations at re-employment among all workers who had unemployment spells which lasted at least x months.

After applying the **Γ**-correction discussed in the previous section, Figure 1a shows that 44.4% of workers who had at least one month in unemployment changed occupation at re-employment, while 53.7% of workers who had at least 9 months in unemployment changed occupation at re-employment. A similar pattern arises when considering mobility across four task-based occupational categories: non-routine cognitive, routine cognitive, non-routine manual and routine manual occupations (NRMC). In this case, the average mobility rate is still about 30% for workers who had at least one month in unemployment, and about 38% among those who spent at least 9 months in unemployment. This evidence thus shows that gross occupational mobility at re-employment is *high* and *increases moderately* with unemployment duration. The moderate increase refers to the fact that a large proportion of long-term unemployed, over 45%, still return to their previous occupation at re-employment.[9]

[9] Kambourov and Manovskii (2008) compare two measures of year-to-year occupational mobility of pooled employer movers and stayers using the PSID, one which includes and one that excludes the unemployed. They find that the inclusion of unemployed workers raises the year-to-year occupational mobility rate by 2.5 percentage points, using a two-digit aggregation. In Appendix A and the Supplementary Appendix E we relate in more detail our analysis to theirs. Moscarini

Table 1: Occupational mobility and unemployment duration

	Γ-corrected mobility			Alternative measures		Mobility, uncorrected		
	(i) U	(ii) NUN	(iii) U,NRMC	(iv) U,Occ*Ind	(v) Retro	(vi) U	(vii) U	(viii) U
Ave mob.	0.444***	0.467***	0.296***	0.389***	0.446***	0.531***	0.531***	0.531***
	(0.006)	(0.004)	(0.003)	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)
OLS dur coef.	0.017***	0.018***	0.012***	0.012***	0.011***	0.014***	0.015***	0.014***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Controls	-	-	-	D, T	T	-	D,T	D, T, S.O.
Obs	19,251	27,833	19,277	18,983	5,410	19,115	19,115	19,115

*/**/*** denotes significance at 10%, 5%, 1% level. Standard Errors on the coefficient on unemployment duration in months (dur. coef) in the regressions on the corrected data are calculated by clustering at the months of completed duration; in the raw SIPP data, by taking the multi-stage stratified survey design into account. Underlying the regression sample are spells with completed durations between 1 and 14 months, not involving agricultural occupations; for further restrictions, see data construction appendix. Underlying sample of average rates contains completed durations between 1 and 18 months, and is further restricted to those workers who are at least 14 months consecutively in sample before being hired, to avoid left-censoring issues. ***Controls***: D=demographic controls (gender, race, education, and a quartic in age); T=time controls (linear time trend, and a dummy for the classification in which data was originally reported); S.O.= source occupation.

Figure 1a also shows that the uncorrected gross mobility rates are only about 9 percentage points above the Γ-corrected ones. This adjustment stands in contrast with the much larger one needed when considering independently-interviewed pooled samples of employer movers and stayers. Indeed, we find that the average occupational mobility rate among all workers is around 27% (when measured at yearly intervals) under independent interviewing, but only about 40% of these observed occupational changes are genuine. The discrepancy arises as our estimated miscoding propensity implies that there is a 20% chance of true stayers appearing as movers and hence a large proportion of the observed mobility rate among all workers can be explain through spurious changes. This leaves relatively little room for true movers (who face a much small probability of being classified spuriously as stayers) to cover the remainder 7%. In contrast, the same miscoding propensity alone will not be able to generate a mobility rate among the unemployed that is close to the 54% observed in Figure 1a. Instead, a large amount of true mobility is required, whereby the larger presence of true movers implies that the proportion of true occupational stayer at risk of being spurious movers is smaller. Ultimately, this implies a smaller absolute adjustment (from 54% to 44%) in the mobility rate of the unemployed relative to all workers. In relative terms this difference is even more substantial (see Appendix A).

Table 1 shows that the high level of mobility and moderate loss of attachment with duration depicted in Figure 1a is robust to alternative specifications. Columns (i) and (iii) summarise the profiles depicted in this figure. Column (ii) finds a very similar pattern when considering non-employment spell that include at least one month of unemployment (NUN-spells). Column (iv) instead uses simultaneous mobility across major occupational and industrial groups, generating a very similar profile relative to (i). This measure is considered less sensitive to miscoding as it typically requires errors to be made simultaneously along two dimensions. Column (v) uses self-reported occupational mobility across non-employment spells, constructed from the end date of the previous job, start date of the current job and self-reported duration of the 'kind of work' the worker is currently performing. This measure captures the worker's own perception of occupational mobility and is not based occupational

and Thomsson (2007) find high occupational mobility among employer-to-employer movers in the CPS, using a sample of workers who changed employers directly or with an intervening spell of non-employment of at most one month.

coding. Here we also find a very similar profile to that of (i) in terms of levels and slope.

Column (vi) reports the estimates based on the uncorrected SIPP data. Compared to (i), mobility in (vi) is not only higher but the coefficient on duration is lower. Miscoding provides an explanation. Since in short unemployment spells true occupational staying is more common, miscoding creates relatively more spurious mobility and therefore our method corrects more the short spells, leading to a steeper duration coefficient in (i).

Demographics Column (viii) shows that controlling for gender, race or education does not change the moderate loss of attachment to an occupation as unemployment duration increases. Indeed, in the Supplementary Appendix we show that the high occupational mobility and the moderate loss of attachment with duration is shared across men and women, high-school and college graduates. However, we find that the $\mathbf{\Gamma}$-corrected level of gross occupational mobility through unemployment decreases substantially with age, from 52.5% when young, (20-30yo), to 39.7% when prime-aged, (35-55yo). Figure 1b shows that the mobility-duration profile of prime-aged workers is below of that of young workers typically by about 9-13 percentage points but has a very similar slope. This is suggestive of an increased attachment of prime-aged workers to their occupation at any duration.

Mobility by occupation Figure 2 shows that most occupations share high $\mathbf{\Gamma}$-corrected mobility rates. Here the total number of unemployment spells corresponds to the rectangular area between zero and 100 in the x-axis and zero and one on the y-axis. The height of each (light-colored) bar shows the ($\mathbf{\Gamma}$-corrected) proportion of workers, given employment in a particular previous occupation, who changed occupations at re-employment. The width of each bar corresponds to the proportion (in percentage points) of each occupation in the inflow into unemployment. Occupations are then sorted in decreasing order by workers' gross mobility. Those occupations within the same task-based category are displayed in the same color. We find that occupations with gross occupational mobility rates above 40% cover more than 80% of all unemployment spells. Apart from small and specialized occupations (as engineers, architects, and doctors), construction is the only large occupation with a meaningfully lower occupational mobility across unemployment spells, which is still close to 25%.

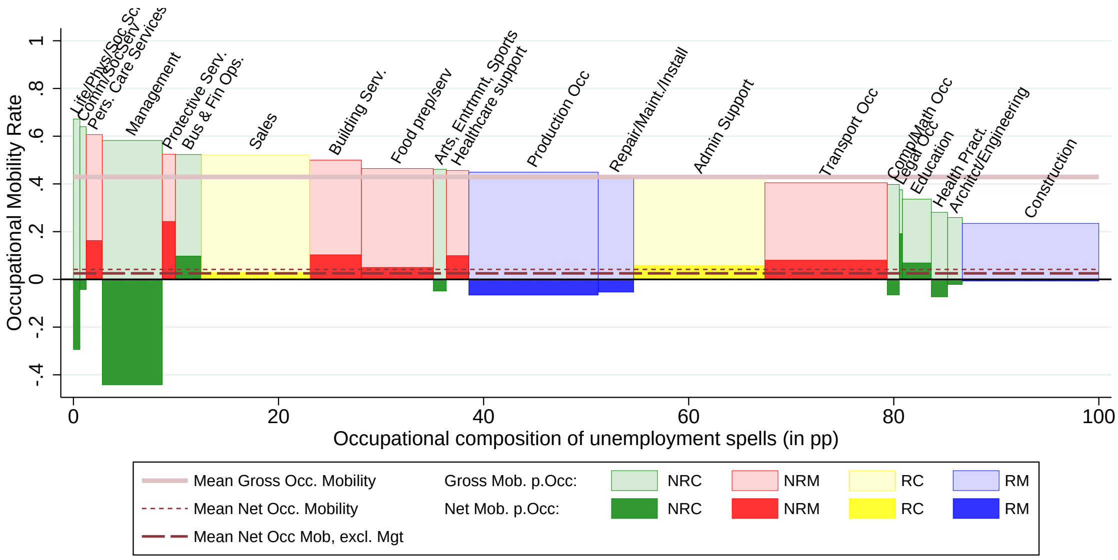
The moderate increase of occupational mobility with unemployment duration is also shared across (source) occupations. In particular, column (ix) in Table 1 shows that controlling for occupational fixed effects does not meaningfully affect the estimated duration coefficient. Moreover, we find that one cannot reject the equality of the occupation-specific slopes (and semi-elasticities) across all occupations. This suggests that the slope of the mobility-duration profile does not arise because some occupations with relatively high unemployment durations have relatively high occupational outflows.

In summary, our evidence presents a clear picture. Gross occupational mobility among the unemployed is high and increases moderately with duration. This pattern is robust to several alternative specifications and widely shared across demographic groups and occupations.

2.3 Excess and Net Mobility

To assess the importance of occupational moves that result in certain occupations experiencing net inflows (outflows) through unemployment, we now divide gross occupational mobility into net and

Figure 2: Gross and Net Occupational Mobility per Occupation



excess mobility. The dark bars in Figure 2 depict the $\mathbf{\Gamma}$-corrected net mobility flows per occupation. The height of each bar corresponds to the proportion of the unemployment spells that originate from a given occupation that are needed to cover the total net flows from that occupation. A positive value refers to net inflows, while a negative value refers to net outflows. The area of each dark colored bar gives the occupation-specific net flows as a proportion of all unemployment spells.

It is evident that net flows are an order of magnitude smaller than gross flows across the vast majority of occupations. The proportion of all unemployment spells that are needed to cover total net flows is just 4.2%, while the proportion of all unemployment spells that are needed to cover total gross flows is 44.4%. The proportion of all gross occupational flows that are necessary to generate the observed net flows between major occupations is then 9.5%. The remaining 90.5% of gross flows represent excess mobility; that is, occupational moves that offset each other.[10] These results are robust to alternative classifications and considering NUN spells instead of only unemployment spells.

Although the $\mathbf{\Gamma}$-correction *lowers* excess mobility across major occupational groups by around 10 percentage points, a 20% decrease, it also *raises* observed net flows. The proportion of unemployment spells that are needed to cover all net mobility rises from 3.6% (uncorrected) to 4.2% (corrected), a nearly 15% increase. To understand why this arises, consider the true net mobility transition flow matrix, $\mathbf{M_{net}}$, and note that this matrix does not have mass on its diagonal. Under independent interviewing coding errors imply that the true net mobility matrix would be observed as $\mathbf{M^I_{net}} = \mathbf{\Gamma}'\mathbf{M_{net}}\mathbf{\Gamma}$, which could have mass on the diagonal and hence biasing downwards net mobility flows.

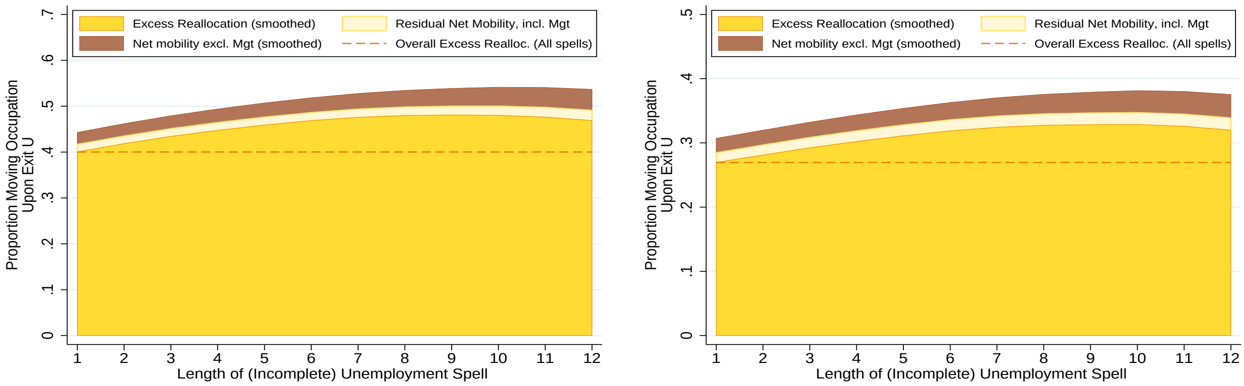
Even though net mobility is small, we find clear patterns among the net flows across task-based categories. In particular, more unemployed workers have left jobs across all routine manual occupations than will find jobs in these occupations, while the opposite is true for non-routine manual

[10] Using pooled samples of employer movers and stayers, Jovanovic and Moffitt (1990) and Kambourov and Manovskii (2008) have also highlighted the importance of excess relative to net mobility across industries or occupations. Net mobility is calculated as the absolute value of the difference between inflows and outflows per occupation, summed up over all occupations, *and divided by two* as one person net inflow in some occupation is also counted as a net outflow some other occupation.

Figure 3: Gross, Net and Excess Occupational Mobility by unemployment duration

(a) Major Occupational Groups (2000 SOC)

(b) Task-based categories (NRC/RC/NRM/RM)



occupations. These patterns are consistent with the job polarization literature which documents the decline of routine manual jobs (e.g. Autor et al. 2003 and Cortes et al., 2020). We also find that non-routine cognitive occupations experience a net outflow. This latter pattern is dominated by former managers, who typically do not return to this occupation when regaining employment.[11]

Excess and net mobility-duration profile To what extent excess and net mobility account for the mobility-duration profile described above and hence for the degree of attachment of workers to occupations? Figure 3 displays the decomposition of the Γ-corrected mobility-duration profile depicted in Figure 1a into three categories: excess mobility, net mobility among non-management occupations (by dropping all the management flows), and the difference between these two, which we label "residual net mobility". The horizontal dashed line across these graphs reflect the average excess mobility rate among those workers who had at least one month in unemployment. We find that excess and net mobility increase with unemployment duration in both the major occupational and task-based categories. However, the increase of net mobility is still small relative to the increase in excess mobility.

2.4 Repeat Mobility

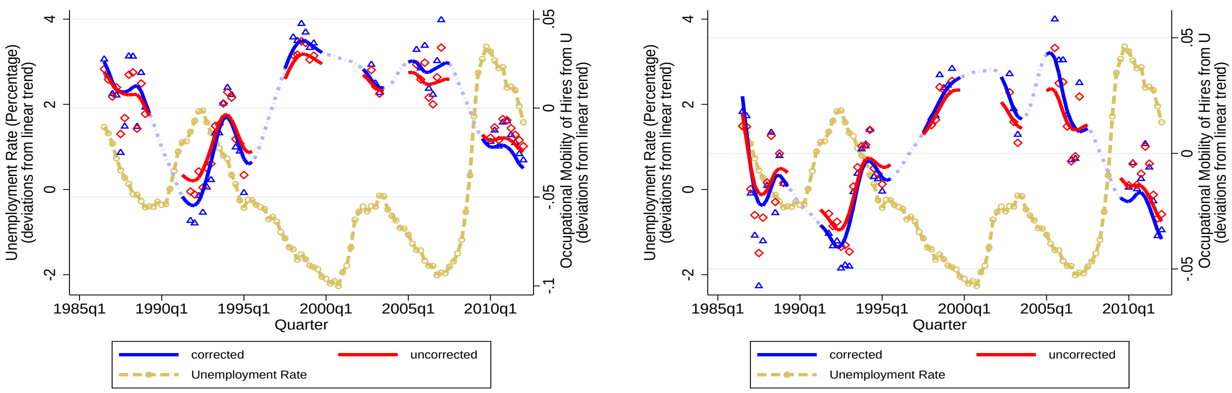
The SIPP allows us to investigate whether workers who changed (did not change) occupations after an unemployment spell, will change occupation subsequently, in a following unemployment spell. These "repeat mobility" statistics provide information about workers' attachment to occupations across unemployment spells.[12] Here we can also use the Γ-correction to counteract coding error in three-

[11] If we exclude managers from our sample, only 2.1% (instead of 4.2%) of unemployment spells are needed to cover the net mobility flows, which is 7.4% (instead of 9.5%) of all grows flows across the reminding occupations.

[12] Our repeat mobility statistics are measured within the SIPP 3.5 to 5 years windows and are based on 610 of observations of individuals with multiple spells across all panels when considering only pure unemployment spells and 1,306 when considering non-employment spells that include months of unemployment. For further details see Supplementary Appendix G. Note that workers with two consecutive unemployment spells within this window are not necessarily representative of all unemployed workers, nor of behavior in unemployment spells that are further apart. Nevertheless, these statistics are valuable and will inform our modelling choices and quantitative analysis, where we take care to construct our simulated measures in the same way as we do in the SIPP.

Figure 4: Cyclicality of occupational mobility

(a) 2000 Census Occupation Classification (b) 4 Task-based occupation super-categories



occupation histories (surrounding two unemployment spells).[13] This is important, as coding error does not only affect mobility measurement directly, but also the division of observations into 'movers' and 'stayers', based on the first unemployment spell.

After correcting for coding errors we find that from all those stayers who became unemployed once again, 64.9% of these workers remain in the same occupation after concluding their second unemployment spell. This percentage is higher for prime-aged workers, 69.3%, and lower for young workers, 57.1%. However, the loss of occupational attachment itself also persists. Among workers who re-enter unemployment after changing occupations in the preceding unemployment spell, we find that 55.8% of these workers move again. This percentage is lower for prime-aged workers, 50.8%, and higher for the young, 63.8%.[14] Supplementary Appendix E shows a similar pattern in the PSID.

2.5 Occupational Mobility of the Unemployed over the Cycle

Unemployed workers' attachment to their previous occupations changes over the business cycle. We find that in expansions unemployed workers change occupations more frequently at re-employment than in recessions. Figure 4a depicts the linearly de-trended time series of the $\mathbf{\Gamma}$-corrected and uncorrected gross mobility rates across major occupational groups, together with the linearly de-trended unemployment rate.[15] Both occupational mobility series are procyclical. The correlation between the unemployment rate and the $\mathbf{\Gamma}$-corrected occupational mobility rate is -0.62, while with the uncor-

[13] With O the total number of occupations, let the matrix $\mathbf{M^r}$ (with elements m^r_{ijk}) be the $O \times O \times O$ matrix of true repeat flows. Then, this matrix relates to the *observed* repeat flow matrix $\mathbf{M^{r,obs}}$ through $\mathbf{vec}(\mathbf{M^r})' = \mathbf{vec}(\mathbf{M^{r,obs}})'(\mathbf{\Gamma} \otimes \mathbf{\Gamma} \otimes \mathbf{\Gamma})^{-1}$, where $\mathbf{vec}(\mathbf{M})$ is the vectorization of matrix $\mathbf{M}$, and $\otimes$ denotes the Kronecker product. Since $\mathbf{\Gamma}$ is invertible, $\mathbf{\Gamma} \otimes \mathbf{\Gamma} \otimes \mathbf{\Gamma}$ is also invertible.

[14] In the uncorrected data the probability of an occupational stay after an occupational stay for all, prime-aged and young workers is 59.0%, 63.4% and 51.6%, respectively; while the probability of an occupational move after an occupational move is 62.0%, 58.5% and 67.8% for all, prime-aged and young workers.

[15] The dashed segments of the time series show were we have interpolated the data because of non-overlapping SIPP panels, combined with our sampling restrictions (to avoid censoring issues). Note that for none of the reported statistics we use the interpolated data. See Supplementary Appendix C for details. There we also show that this negative correlation is robust to different methods of filtering/de-trending, measurement at time of separation or at re-hiring, robust to controlling for observables and shared across mobility measures based on different classifications.

Table 2: Occupational mobility and unemployment duration over the business cycle

	Γ-corrected mobility			Alternative measures		Mobility, uncorrected		
	(i) U	(ii) NUN	(iii) U,NRMC	(iv) U,Occ*Ind	(v) Retro	(vi) U	(vii) U	(viii) U
Panel A: Mobility regression, not controlling for non-employment duration								
HP U	-0.154**	-0.109**	-0.088*	-0.114**	-0.765***	-0.114**	-0.139***	-0.129***
	(0.062)	(0.048)	(0.049)	(0.047)	(0.176)	(0.049)	(0.045)	(0.043)
Panel B: Mobility regression, controlling for non-employment duration								
HP U	-0.199***	-0.140***	-0.128**	-0.155***	-0.877***	-0.150***	-0.188***	-0.174***
	(0.063)	(0.047)	(0.049)	(0.046)	(0.175)	(0.050)	(0.045)	(0.044)
Dur coef	0.0161***	0.0165***	0.0125***	0.0112***	0.015**	0.0133***	0.0153***	0.0142***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)	(0.002)
Controls	-	-	-	D,T	D,T	-	D,T	D,T, S.O.

*/**/*** denotes significance at 10%, 5%, 1% level. See the Supplementary Appendix for further description. Sample is restricted to quarters where the data allows the full spectrum of durations between 1-12 months to be measured.

rected series the correlation is -0.63. Figure 4b shows that the same cyclical pattern holds across the four task-based categories, with a correlation of -0.47 with the Γ-corrected series and of -0.49 with the uncorrected series. We find that this procyclicality is also present when using linearly de-trended log productivity instead of the unemployment rate.[16]

Panel A of Table 2 displays the results of regressing workers' occupational mobility on HP-filtered unemployment rates without controlling for unemployment (non-employment) duration. We find that occupational mobility among the unemployed is procyclical for both the Γ-corrected and uncorrected series. Note the larger values (in absolute terms) of the unemployment coefficients when using the Γ-corrected series. This arises as coding errors will generate more spurious mobility in times were there are more true stayers. Further, controlling for demographic characteristics and the identity of source (and destination) occupation does not meaningfully affect the unemployment coefficient. This indicates that the procyclicality of occupational mobility is not the result of a compositional shift towards occupations or demographics characteristics which are associated with higher mobility when the economy is in an expansion. Instead this procyclicality seems to be shared across a wide range of occupations and demographic groups. A similar result is also found when using the PSID.

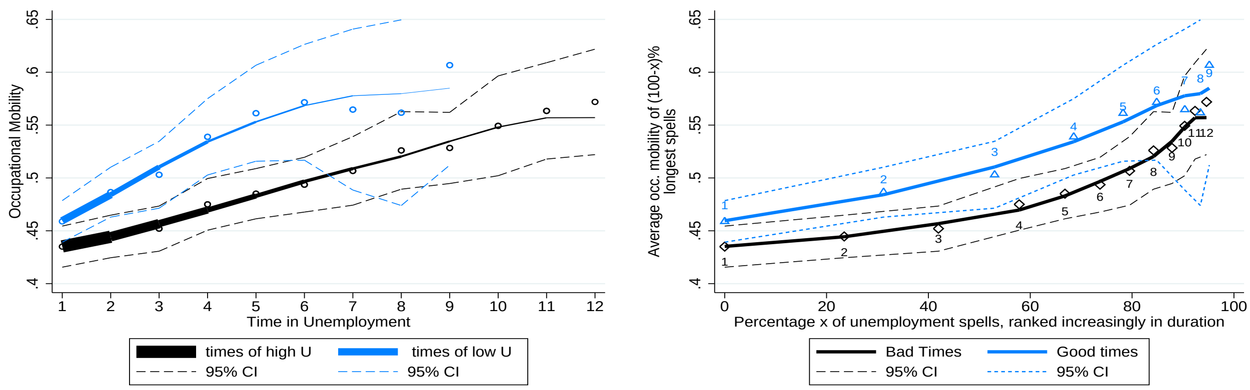
Panel B displays the results from regressing workers' occupational mobility on HP-filtered (log) unemployment rates, controlling for unemployment duration. Here we find an even more responsive mobility rate to changes in unemployment, as now the unemployment coefficient only captures the vertical shift of the mobility-duration profile. In Panel A, however, the unemployment coefficient captures a downward shift and a simultaneous rightward-and-upward movement along the mobility-duration profile, as unemployment spells are longer in recessions.

The cyclicality of the mobility-duration profile Figure 5a plots the mobility-duration profile for those spells that ended in times of high unemployment and the profile for those spells that ended in times of low unemployment. Times of high (low) unemployment are defined as periods in which the de-trended (log) unemployment rate was within the bottom (top) third of the de-trended (log)

[16]The procyclicality of occupational mobility among the unemployed coincides with that of pooled samples of employer movers and stayers (see Kambourov and Manovskii, 2008, and Moscarini and Thomsson, 2007), even tough those who changed employers through unemployment represent a small part of all employed workers.

Figure 5: Gross Occupational Mobility - Unemployment Duration Shift over the Cycle

(a) Occ. Mobility - Duration Profile Shift (b) Occ. Mobility - Duration Distribution Shift



unemployment distribution. The thickness of the profiles indicates the amount of spells surviving at a given duration, showing the faster reduction of spells with duration in expansions. This graph shows that occupational mobility at any unemployment duration is lower in recessions, and that both in times of high and low unemployment, relatively longer-term unemployment is associated with a moderate loss of attachment to the previous occupation.

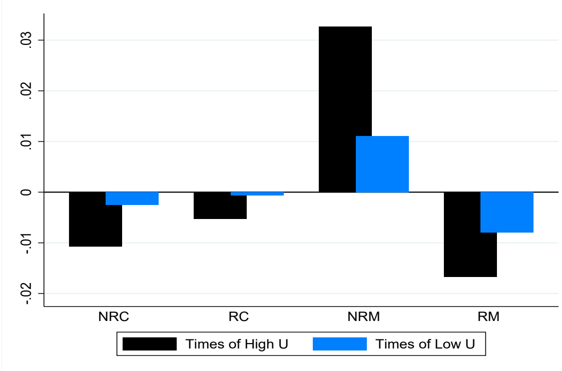
Figure 5b shows that the effect of downturns on the mobility-duration profile is not simply the result of most spells lasting longer. This figure compares the Γ-corrected occupational mobility rates associated with a given rank in the distribution of unemployment spells across times of high and low unemployment. The horizontal axis indicates the rank of unemployment spells by duration, while the vertical axis indicates the occupational mobility associated with the unemployment spells of at least rank x. We observe that in recessions mobility is lower at any percentile of the duration distribution and statistically significantly so up until we confine ourselves to about the 20% longest spells.

The cyclicality of net occupational mobility Figure 6 shows the cyclical behavior of the Γ-corrected net mobility rates for each of the four task-based categories. It is clear that net mobility increases in periods of high unemployment relative to periods of low unemployment.[17] Note that the job polarization pattern described earlier remains across the business cycle. Both in expansions and downturns the routine manual occupations lose more workers than workers get re-employed in these occupations. In contrast, the non-routine manual occupations exhibit net inflows of workers in both expansions and downturns.[18] The countercyclicality of net mobility therefore implies that the cyclicality of excess mobility is the main driver behind the procyclical behaviour of gross occupational mobility among unemployed workers.

[17] In this case we define times of high (low) unemployment as periods in which the de-trended (log) unemployment rate was within the top (bottom) third (half) of the de-trended (log) unemployment distribution. We chose this partition as it minimises small sample bias. In Supplementary Appendix C we show that the same patterns hold when defining periods of low and high unemployment in many different ways.

[18] Kambourov and Manovskii (2008) find countercyclical net mobility and procyclical gross mobility among a pooled sample of employer stayers and movers.

Figure 6: Net Occupational Mobility - Task-based categories (excl. Managers)



2.6 Unemployment duration of occupational movers

Averaged over the business cycle, movers take 0.5 months longer in unemployment than stayers, also after taking account of demographic characteristics and occupational identities. This difference is economically significant: it represents nearly half of the differences between the average unemployment spell in periods of high versus low unemployment. In recession, the difference grows to 1.11 months. This increase does not result from cyclically different demographics of unemployed movers or because they are more likely to be in long-duration occupations in recessions (see Supplementary Appendix D for a formal regression analysis). Although the occupational mobility of the unemployed decreases in a recession, the lengthening of unemployment spells among movers is proportionally stronger. Thus, occupational movers contribute to the increase in aggregate unemployment, and especially to the increase in long-term unemployment.

3 Theoretical Framework

We now develop a theory of occupational mobility of the unemployed to explain the above empirical results. Our model incorporates aggregate uncertainty, occupational-wide productivity differences and time-variant worker heterogeneity.

3.1 Environment

Time is discrete $t = 0, 1, 2, \ldots$ A mass of infinitely-lived, risk-neutral workers is distributed over a finite number of occupations $o = 1, \ldots, O$. At any time t, workers within a given occupation differ in two components: an idiosyncratic productivity, z_t, and human capital, x_h. We interpret the z-productivity as a "career match" which captures in a reduced form the changing career prospects workers have in the occupation they are currently attached to (see Neal, 1999). Workers' z-productivities follow a common and exogenous first-order stationary Markov process, with transition law $F(z_{t+1}|z_t)$ and $z_t, z_{t+1} \in [\underline{z}, \overline{z}]$, $\underline{z} > 0$ and $\overline{z} < \infty$. The z-productivity realizations affect a worker both in employment and in unemployment and will drive excess occupational mobility in our model. The assumption that the z-productivity process is common across workers and occupations is motivated by our evidence showing that the change of occupational mobility with unemployment duration does not seem to differ across occupations or demographic groups.

To capture the different levels of attachment to occupations found across age groups, we assume workers' accumulate stochastically occupational human capital through a learning-by-doing process. In period t an employed worker with human capital level x_h increases his human capital to x_{h+1} with probability $\chi^e(x_{h+1}|x_h)$, where $\chi^e(x_{h+1}|x_h) = 1 - \chi^e(x_h|x_h)$, $x_h < x_{h+1}$, $h = 1, \ldots, H$ and $x_H < \infty$. A worker's human capital may also depreciate stochastic with unemployment. An unemployed worker with human capital level x_h decreases his human capital to x_{h-1} with probability $\chi^u(x_{h-1}|x_h)$, where $\chi^u(x_{h-1}|x_h) = 1 - \chi^u(x_h|x_h)$, $x_h > x_{h-1}$, $h = 1, \ldots, H$ and $x_H < \infty$.

We model the business cycle through fluctuations in the economy-wide productivity and let A_t denote this aggregate productivity. We assume A_t follows a first-order stationary Markov process with $A_t \in [\underline{A}, \overline{A}]$, $\underline{A} > 0$ and $\overline{A} < \infty$. To generate net occupational mobility we allow some occupations to be more attractive than others in terms of their occupation-wide productivities. Let $p_{o,t}$ denote the occupation-wide productivity of occupation o at time t. We assume $p_{o,t}$ follows a stationary first-order Markov process, with $p_{o,t} \in [\underline{p}_o, \overline{p}_o]$, $\underline{p}_o > 0$ and $\overline{p}_o < \infty$. In principle this process could depend on aggregate productivity A_t. Let $\mathcal{P}_{O,t} = \{p_{o,t}\}_{o=1}^{O}$ denote the vector that contains all the occupation-wide productivities at time t.

There is also a mass of infinitely-lived risk-neutral firms distributed across occupations. All firms are identical and operate under a constant return to scale technology, using labor as the only input. Each firm consists of only one job that can be either vacant or filled. The output of a worker with current productivity z_t and human capital x_h employed in a firm in occupation o is given by the production function $y(A_t, p_{o,t}, z_t, x_h)$. The production function is strictly increasing and continuous in all of its arguments and differentiable in the first three. We assume that any unemployed worker receives b each period. Wages will be determined below.

All agents discount the future at rate β. Workers retire stochastically and afterwards receive a fixed utility flow normalized to zero. They are replaced by inexperienced workers entering the labor force. We can rescale β to incorporate this risk for both workers and firms.

Searching within an occupation A key assumption is that workers with different pairs (z, x_h) do not congest each other in the matching process. This assumption allows us to study business cycle behavior in a tractable way. As a result, an occupation is segmented into many labor markets, one for each pair (z, x_h). Each labor market (z, x_h) has the DMP structure. A constant returns to scale matching function governs the meetings of unemployed workers and vacancies within each market. We assume that all these labor markets have the same random matching technology. Each labor market exhibits free entry of firms, where posting a vacancy costs k per period. When the z-productivity of an unemployed worker with human capital x_h changes to $\hat{z}$, the relevant labor market for this worker becomes $(\hat{z}, x_h)$. Similarly, when the human capital of an unemployed worker changes, the relevant labor market becomes (z, x_{h-1}). Should an employed worker become unemployed the relevant labor market would be the one associated with his current z-productivity and human capital level x_h. Match break-up can occur with an exogenous (and constant) probability δ, but can also occur if the worker and the firm decide to do so, and after a retirement shock. Once the match is broken, the firm decides to reopen the vacancy and, unless retired, the worker stays unemployed until the end of the period.

Searching across occupations Instead of searching for jobs in their own occupation, unemployed workers can decide to search for jobs in different occupations. This comes at a per-period cost c and entails re-drawing their z-productivity process. Workers rationally expect their initial career match in any occupation to be a draw from some distribution $F(z)$, which we take to be the ergodic distribution associated with the aforementioned Markov process $F(z_{t+1}|z_t)$. The i.i.d. nature of the re-draws allows us to capture that some occupational movers end up changing occupations again after a subsequent jobless spell, as suggested by the repeat mobility patterns documented earlier.

Given differences in occupation-wide labor market conditions $p_{o,t}$, workers are not indifferent from which occupation they draw the new z-productivity. To capture that in the data excess mobility is much larger than net mobility, we model search across occupations in the spirit of Fallick (1993). In particular, an unemployed worker who is considering moving away from his current occupation has a limited amount of time to investigate his employment prospects in different occupations. The worker has to decide what proportion $s_{\tilde{o}}$ of his time to devote to obtain the z-productivity from occupation $\tilde{o}$. Let $\mathcal{S}$ denote a vector of $s_{\tilde{o}}$ for all $\tilde{o} \in O^-$, where O^- denotes the set of remaining occupations such that $\sum_{\tilde{o} \in O^-} s_{\tilde{o}} = 1$. The probability that a worker, currently in occupation o, then receives the new z from occupation $\tilde{o}$ is given by $\alpha(s_{\tilde{o}}; o)$, where $\alpha(.; o)$ is a continuous, weakly increasing and weakly concave function with $\alpha(0; o) = 0$ and $\sum_{\tilde{o} \in O^-} \alpha(s_{\tilde{o}}; o) \leq 1$ for all $o \in O$. The latter summation allows for the possibility of not receiving a z-productivity, in which case the process is repeated the following period.

After observing the new z-productivity, the worker must sit out one period unemployed before deciding whether to sample another z-productivity from a different occupation. If the worker decides to sample the above process is repeated. If the worker decides to 'accept' the z-productivity he starts with human capital x_1 in the new occupation. The worker's z-productivity and occupational human capital then evolve as described above.[19]

Timing and state space The timing of the events is summarised as follows. At the beginning of the period the new values of A, $\mathcal{P}_O$, z and x are realised. After these realisations, the period is subdivided into four stages: separation, reallocation, search and matching, and production. Let $\mathcal{G}$ denote the joint productivity distribution of unemployed and employed workers over all occupations. Let $\mathcal{G}^j$ denote this distribution at the beginning of stage j. The state space for a worker currently characterised by (z, x) at the beginning of stage j is described by the vector $\Omega^j = \{z, x, o, A, \mathcal{P}_O, \mathcal{G}^j\}$, and his employment status.

We can show that the equilibrium decision rules have a relevant state space described solely by

[19]Our data suggests that c and the loss of x_h when changing occupation should be incorporated in our model as mobility costs. This is because we find (i) a substantial proportion of stayers among young workers, which are typically associated with low levels of human capital, and (ii) substantial occupational staying among those who moved occupations but subsequently have become unemployed again. Since this occurs within the duration of a SIPP panel, these workers' occupational tenure is low, yet they also display significant occupational attachment.

$\omega = \{z, x, o, A, \mathcal{P}_O\}$, and workers' employment status. To keep notation complexity to a minimum, we present the agents' decision problems and the laws of motion of unemployed and employed workers using this state space. To further simplify notation we leave implicit the time subscripts, denoting the following period with a prime.

3.2 Agents' Decisions

Worker's Problem Consider an unemployed worker currently characterised by (z, x, o). The value function of this worker at the beginning of the production stage is given by

$$W^U(\omega) = b + \beta \mathbb{E}_{\omega'}\left[\max_{\rho(\omega')}\left\{\rho(\omega')R(\omega') + (1-\rho(\omega'))\left[\lambda(\theta(\omega'))W^E(\omega') + (1-\lambda(\theta(\omega')))W^U(\omega')\right]\right\}\right], \tag{1}$$

where $\theta(\omega)$ denotes the ratio between vacancies and unemployed workers currently in labor market (z, x) of occupation o, with $\lambda(.)$ the associated job finding probability. The value of unemployment consists of the flow benefit of unemployment b, plus the discounted expected value of being unemployed at the beginning of next period's reallocation stage, where $\rho(\omega)$ takes the value of one when the worker decides to search across occupations and zero otherwise. The term $R(\omega)$ denotes the expected net value of searching across occupations and is given by

$$R(\omega) = \max_{\mathcal{S}(\omega)}\Big(\sum_{\tilde{o}\in O^-}\alpha(s_{\tilde{o}}(\omega))\int_{\underline{z}}^{\overline{z}} W^U(\tilde{z}, x_1, \tilde{o}, A, \mathcal{P}_O)dF(\tilde{z}) + (1 - \sum_{\tilde{o}\in O^-}\alpha(s_{\tilde{o}}(\omega)))\hat{W}^U(\omega) - c\Big), \tag{2}$$

where $\hat{W}^U(\omega) = b + \beta \mathbb{E}_{\omega'} R(\omega')$ and the maximization is subject to $s_{\tilde{o}} \in [0, 1]$ and $\sum_{\tilde{o}\in O^-} s_{\tilde{o}} = 1$. The formulation of $\hat{W}^U(\omega)$ is helpful as it implies that $R(\omega)$ and $\{s_{\tilde{o}}\}$ become independent of z. It is through $R(\omega)$ that expected labor market conditions in other occupations affect the value of unemployment, and indirectly the value of employment, in the worker's current occupation. The worker's decision to reallocate is captured by the choice between the expected net gains from drawing a new $\tilde{z}$ in another occupation and the expected payoff of remaining in the current occupation. The latter is given by the expression within the inner squared brackets in equation (1).

Now consider an employed worker currently characterised by the pair (z, x) in occupation o. The expected value of employment at the beginning of the production stage, given wage $w(\omega)$, is

$$W^E(\omega) = w(\omega) + \beta \mathbb{E}_{\omega'}\Big[\max_{d(\omega')}\{(1 - d(\omega'))W^E(\omega') + d(\omega')W^U(\omega')\}\Big]. \tag{3}$$

The second term describes the worker's option to quit into unemployment in next period's separation stage. The job separation decision is summarised in $d(\omega')$, such that it take the value of δ when $W^E(\omega') \geq W^U(\omega')$ and the value of one otherwise.

Firm's Problem Consider a firm posting a vacancy in labor market (z, x) in occupation o at the start of the search and matching stage. The expected value of a vacancy solves the entry equation

$$V(\omega) = -k + q(\theta(\omega))J(\omega), \tag{4}$$

where $q(.)$ denotes firms' probability of finding an unemployed worker and $J(\omega)$ denotes the expected value of a filled job. Free entry implies that $V(\omega) = 0$ for all those sub-markets that yield a $\theta(\omega) > 0$, and $V(\omega) \leq 0$ for all those sub-markets that yield a $\theta(\omega) \leq 0$. In the former case, the entry condition simplifies (4) to $k = q(\theta(\omega))J(\omega)$.

Now consider a firm employing a worker currently characterized by the pair (z, x) at wage $w(\omega)$ in occupation o. The expected lifetime discounted profit of this firm at the beginning of the production stage can be described recursively as

$$J(\omega) \;=\; y(A, p_o, z, x) - w(\omega) + \beta \mathbb{E}_{\omega'}\Big[\max_{\sigma(\omega')}\Big\{(1-\sigma(\omega'))J(\omega') + \sigma(\omega')V(\omega')\Big\}\Big], \tag{5}$$

where $\sigma(\omega')$ takes the value of δ when $J(\omega') \geq V(\omega')$ and the value of one otherwise.

Wages We assume that wages are determined by Nash Bargaining. Consider a firm-worker match currently associated with the pair (z, x) in occupation o such that it generates a positive surplus. Nash Bargaining implies that the wage, $w(\omega)$, solves

$$(1-\zeta)\Big(W^E(\omega) - W^U(\omega)\Big) = \zeta\Big(J(\omega) - V(\omega)\Big), \tag{6}$$

where $\zeta \in [0, 1]$ denotes the worker's exogenous bargaining power. This guarantees that separation decisions are jointly efficient, $d(\omega) = \sigma(\omega)$.

In what follows we impose a Cobb-Douglas matching function and the Hosios condition, such that $1 - \zeta = \eta$, where η denotes the elasticity of the job finding probability with respect to labor market tightness. This will guarantee that firms post the efficient number of vacancies within labor markets. It will also guarantee efficiency of our decentralized economy.

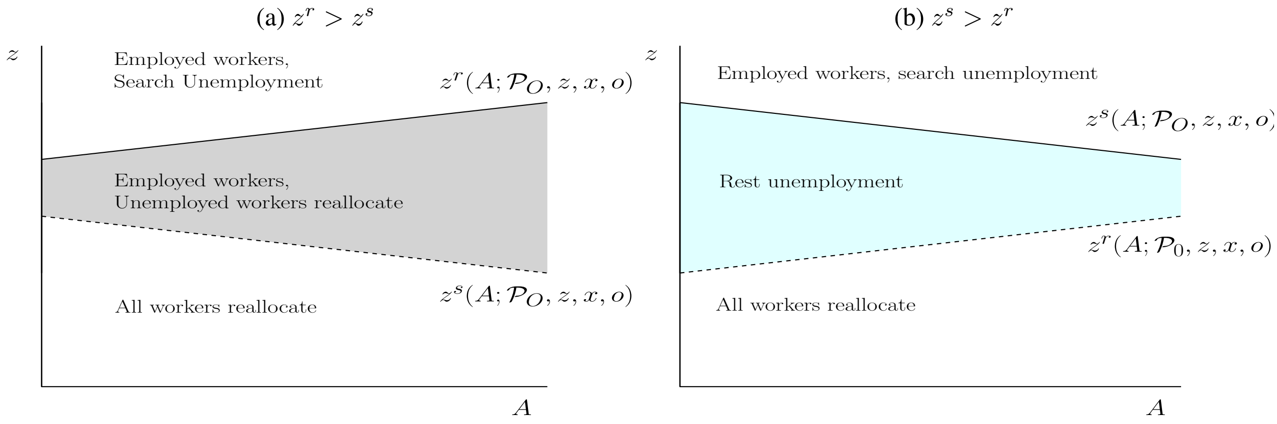
3.3 Equilibrium and Characterization

We focus on equilibria in which the value functions and decisions of workers and firms in any occupation only depend on $\omega = \{z, x, o, A, \mathcal{P}_O\}$ and workers' employment status. In this type of equilibria outcomes can be derived in two steps. In the first step, decision rules are solved independently of the distribution $\mathcal{G}$, using (1)-(5). Once those decision rules are determined, we fully describe the dynamics of $\mathcal{G}$, using the workers' flow equations. This recursive structure follows Menzio and Shi (2010, 2011) and we therefore label it a Block Recursive Equilibrium (BRE). To prove existence and uniqueness we build on their proofs but incorporate the value of reallocation across occupations and show it preserves the block recursive structure. The formal definition of the BRE is relegated to Appendix B, where we also present the derivation of the flow equations and the proofs of all the results of this section.

Existence Let $M(\omega) \equiv W^E(\omega) + J(\omega)$ denote the joint value of the match. To prove existence and uniqueness of the BRE we define an operator T that maps $M(\omega)$, $W^U(\omega)$ and $R(\omega)$ from the appropriate functional space into itself, with a fixed point that implies a BRE. Given this result and the Banach's Fixed Point Theorem, this fixed point exists and is unique.

To prove efficiency we show that the unique solution to the planner's problem in the general state space Ω coincides with the solution to the decentralized economy problem in the space ω. The key step is to ensure that a worker's value of searching across occupations coincides with the planner's value of making the worker search across occupations.

Figure 7: Relative positions of the reservation productivities



Proposition 2. *Given $F(z'|z) < F(z'|\tilde{z})$ for all z, z' when $z > \tilde{z}$: (i) a BRE exists and it is the unique equilibrium; and (ii) the BRE is constrained efficient.*

Characterization The decision to separate from a job and the decision to search across occupations can be characterised by z-productivity cutoffs, which are themselves functions of A, $\mathcal{P}_O$, o and x. The job separation cutoff function, $z^s(.)$, is related to the one found in Mortensen and Pissarides (1994) because it characterises endogenous separations. However, in our setup z refers to the worker's idiosyncratic productivity in an occupation, rather than to a match-specific productivity with a firm. This difference implies that when the worker becomes unemployed, his z-productivity is not lost or is reset when re-entering employment in the same occupation. Instead, the worker's z-productivity continuous evolving during the unemployment spell. It is only when the worker searches across occupations that he can reset his z-productivity. The reallocation cutoff function, $z^r(.)$, determines when an unemployed worker decides to search across occupations. The latter occurs if and only if $z < z^r(.)$.

The relative position and the slopes of $z^r(.)$ and $z^s(.)$ are crucial determinants of the long-run and cyclical outcomes in our model. To show this, we first discuss the implications of their relative position and then of their slopes. Figure 7a illustrates the case in which $z^r > z^s$ for all A, holding constant $\mathcal{P}_O$, o and x. Here having a job makes a crucial difference on whether a worker stays or leaves his occupation. When an employed worker has a current $z \in [z^s, z^r)$, the match surplus is enough to keep him attached to his occupation. For an unemployed worker with a current z in the same interval, however, the probability of finding a job is sufficiently small to make searching across occupations the more attractive option, even though this worker could generate a positive match surplus if he were to become employed in his pre-separation occupation. For values of $z < z^s$, all workers search across occupations. For values of $z \geq z^r$, firms post vacancies and workers remain in their occupations, flowing between unemployment and employment over time as in the canonical DMP model.

Figure 7b instead shows the case in which $z^s > z^r$ for all A. Here workers who endogenously separate into unemployment, at least initially, do not search across occupations, while firms do not create vacancies in labor markets associated with values of $z < z^s$. Workers are *rest unemployed* during the time in which their $z \in [z^r, z^s)$: they face a very low – in the model (starkly) zero –

contemporaneous job finding probability, but still choose to remain attached to their occupations. The stochastic nature of the z-productivity process, however, implies that these workers can face a positive expected job finding probability for the following period. Only after the worker's z-productivity has declined further, such that $z < z^r$, the worker searches for a new z across occupations. For values of $z \geq z^s$, the associated labor markets function as in the canonical DMP model.

An unemployed worker is then considered *search unemployed* during the time in which his $z \geq z^s$, as in the associated labor markets firms are currently posting vacancies. A worker whose current $z < z^r$ is considered *reallocation unemployed* only during the time in which he is trying to find another occupation that offers him a $z > z^r$. Once he finds such an occupation, he continues his unemployment spell potentially with periods in search and rest unemployment, depending on the relative position of z^s and z^r and the initial draw and evolution of his z-productivity in such an occupation. The stochastic nature of the z process implies that search, rest and reallocation unemployment are not fixed characteristics, but transient states during an unemployment spell. Therefore, to be consistent with the analysis of Section 2, an *occupational mover* is a worker who left his old occupation, went through a spell of unemployment (which could encompass all three types of unemployment) and found a job in a different occupation.

A key decision for an unemployed worker is whether to remain in his occupation, waiting for his z-productivity to improve, or to search across occupations, drawing a new z-productivity. Periods of rest unemployment arise when the option value of waiting in unemployment is sufficiently large. However, search frictions imply that there is also an option value associated with waiting in employment in an existing job match. In the face of irreversible match destruction, workers remain employed at lower output levels relative to the frictionless case because of potential future improvements in their z-productivities. This drives the separation cutoff function down.

The tension lies in that these two waiting motives work against each other. Which one dominates depends on parameter values. Using a simplified version of the model without aggregate or occupation-specific shocks, we show that the difference $z^s - z^r$ increases when c, b or x increase (see Appendix B.1). Although it is intuitive that a higher c or x reduces z^r by making occupational mobility more costly, they also reduce z^s by increasing the match surplus and making employed workers less likely to separate. We show that, overall, the first effect dominates. A rise in b decreases z^r by lowering the effective cost of waiting, while decreasing the match surplus by increasing $W^U(.)$ and hence increasing z^s, pushing towards rest unemployment. We also show that a higher degree of persistence in the z-productivity process decreases $z^s - z^r$ as it decreases the option value of waiting.

Figure 7 shows the case of countercyclical job separation decisions ($\partial z^s(.)/\partial A < 0$) and procyclical occupational mobility decisions ($\partial z^r(.)/\partial A > 0$), as suggested by the data. The relative position of z^s and z^r is an important determinant of the cyclicality of occupational mobility decisions. Using a simplified version of the model without occupation-specific shocks, we show that when $z^s > z^r$ we obtain procyclical occupational mobility decisions without the need of complementarities in the production function (see Appendix B.1). This arises as search frictions create endogenous complementarities between the wage and the job finding probability. As aggregate productivity increases, the value of drawing a z increases and more so at the top of $F(z)$, making the decision to reallocate more attractive and steepens $z^r(.)$ relative to the frictionless case. In addition, the presence of rest unemployment reduces the opportunity cost of occupational mobility as it becomes less responsive to aggregate shocks. This occurs because any change in A does not immediately affect the utility flow of rest unemployed workers.

The relative position of z^s and z^r is also an important determinant of the cyclicality of job separation decisions. When z^s is sufficiently above z^r, job separation decisions mainly reflect whether or not an employed worker should wait unemployed in his current occupation for potential improvement of his z. Occupational mobility is only one possible future outcome and hence it is discounted. This implies that a sufficiently large $z^s - z^r > 0$ moderates the feedback of procyclical occupational mobility decisions on the cyclicality of job separation decisions.

As the position and slope of the z^s and z^r cutoffs can only be fully determined through quantitative analysis, we now turn to estimate the model and investigate its resulting cyclical properties.

4 Quantitative Analysis

4.1 Calibration Strategy

We set the model's period to a week and the discount factor $\beta = (1-d)/(1+r)$ is such that the exit probability, d, is chosen to match an average working life of 40 years and r such that β matches an annual real interest rate of 4%. To keep the population constant every worker that leaves the economy is replaced by a new unemployed worker. We target occupational mobility statistics based on the 2000 SOC and aggregate the simulated data to 'major' occupational groups and task-based categories (non-routine cognitive NRC, routine cognitive RN, non-routine manual NRM and routine manual RM) as done in Section 2. Our classification error model (Section 2.1) then allows us to easily correct for aggregate and occupation-specific levels of miscoding by imposing the $\mathbf{\Gamma}$-correction matrix on simulated worker occupational flows at the required level of aggregation.

The worker's production function is assumed multiplicative and given by $y_o = Ap_oxz$ for all $o \in O$, chosen to keep close to a 'Mincerian' formulation. The logarithm of aggregate productivity, $\ln A_t$, follows an AR(1) process with persistence and dispersion parameters ρ_A and σ_A. For a given occupation o, the logarithm of the occupation-wide productivity is given by $\ln p_{o,t} = \ln \overline{p}_o + \epsilon_o \ln A_t$, where $\overline{p}_o$ denotes this occupation's constant productivity level and ϵ_o its loading with respect to changes in aggregate productivity. This formulation implies that different occupations can have different sensitivities to the aggregate shock and hence different relative attractiveness to workers over the business cycle.[20] We consider occupation-wide productivity differences at the level of task-based categories, $O = \{NRC, RC, NRM, RM\}$. All major occupations within a task-based category $o \in O$ then share the same $p_{o,t}$. This approach not only simplifies the computational burden by reducing

[20]The evidence presented in Supplementary Appendix C.3 suggests that our approach is consistent with the observed cyclical behaviour of net occupational flows, where the majority of occupations exhibit a very similar cyclical pattern across several recession/expansion periods.

the state space of the calibrated model, but is also consistent with the evidence presented in Figure 2 showing that within the majority of task-based categories all major occupations' net flows exhibit the same sign. To further simplify we normalize both the employment weighted average of $\overline{p}_o$ and of ϵ_o across $o \in O$ to one.

Occupational human capital is parametrized by a three-level process, $h = 1, 2, 3$. We normalize $x_1 = 1$ and set $\chi^e(x_{h+1}|x_h)$ such that the next level is stochastically acquired after five years on average. With probability $\chi^u(x_{h-1}|x_h) = \gamma_d$ the human capital of an unemployed worker depreciates one level until it reaches x_1. The logarithm of the worker's idiosyncratic productivity within an occupation, $\ln z_t$, is also modelled as an AR(1) process with persistence and dispersion parameters ρ_z and σ_z. We include a normalization parameter $\underline{z}_{norm}$ that moves the entire distribution of z-productivities such that measured economy-wide productivity averages one.

The probability that a worker in a major occupation within task-based category o receives the new z from a different major occupation in task-based $\tilde{o}$ is parametrized as $\alpha(s_{\tilde{o}}; o) = \overline{\alpha}_{o,\tilde{o}}^{(1-\nu)} s_{\tilde{o}}^{\nu}$ for all $o, \tilde{o}$ pairs in $O = \{NRC, RC, NRM, RM\}$ and $s_{\tilde{o}} \in [0, 1]$. The parameter $\nu \in [0, 1]$ governs the responsiveness of the direction of search across occupations that is related to differences in p_o. The parameter $\overline{\alpha}_{o,\tilde{o}}$ is a scaling factor such that $\sum_{\tilde{o} \in O} \overline{\alpha}_{o,\tilde{o}} = 1$. It captures the extent to which an unemployed worker in a major occupation within task-based category o has access to job opportunities in another major occupation in task-based category $\tilde{o}$. Since $\sum_{\tilde{o} \in O} \alpha(s_{\tilde{o}}; o) \leq 1$, this formulation implies that if a worker in o wants to obtain a new z with probability one, he will choose $s_{\tilde{o}} = \overline{\alpha}_{o,\tilde{o}}$ for all $\tilde{o} \in O$. If a worker wants to take into account current occupation-wide productivity differences, he will choose $s_{\tilde{o}} \neq \overline{\alpha}_{o,\tilde{o}}$ for at least some $\tilde{o}$. The cost of doing so is the possibility of not receiving a new z at all (i.e. $\sum_{\tilde{o} \in O} \alpha(s_{\tilde{o}}; o) < 1$) and paying c again the following period. The parameter ν determines whether this cost is large or very small, with higher values of ν leading to lower probabilities of not receiving a new z.

To capture the observed long-run changes in the employment-size distribution across task-based categories, we first externally calibrate the initial size distribution in the simulations to match the one observed in the SIPP in 1984. This results in setting the employment proportions for NRC, RC, NRM, RM to 0.224, 0.292, 0.226 and 0.258, respectively, at the start of the simulation. In addition to the occupational mobility decisions of workers who changed employers through a spell of unemployment, we allow this size distribution to change over time due to the mobility decisions of new labor market participants. Let ψ_o denote the exogenous probability that a new entrant to the economy is allocated to task-based category o such that $\sum_{o \in O} \psi_o = 1$. This worker is then randomly allocated to a major occupation within the drawn task-based category at the point of entry, and is allowed to search across occupations to obtain first employment somewhere else.

To allow for differences in the separation rates across young and prime-age workers that are not due to the interaction between z and x_h, we differentiate the probability of an exogenous job separation between low (x_1) and high human capital (x_2, x_3) workers: δ_L and δ_H. Finally, the matching function within each sub-market in any occupation is given by $m(\theta) = \theta^{\eta}$.

The above parametrization then yields a set of 36 free parameters to estimate, where $[c, \rho_z, \sigma_z, \underline{z}_{norm}]$

govern occupational mobility due to idiosyncratic reasons (excess mobility); $[\overline{p}_o, \epsilon_o, \overline{\alpha}_{o,\tilde{o}}, \nu, \psi_o]$ for all $o, \tilde{o} \in \{NRC, RC, NRM, RM\}$ govern occupational mobility due to occupation-wide productivity differences (net mobility); $[x_2, x_3, \gamma_d, \delta_L, \delta_H]$ govern differences in occupational human capital; and the remainder parameters $[k, b, \eta, \rho_A, \sigma_A]$ are shared with standard DMP calibrations. All these parameters are estimated by minimising the sum of squared distances between a set of 95 independent model simulated moments and their data counterparts. For consistent measurement we generate 'pseudo-SIPP panels' within one hundred time-windows each of 30 year length and follow the same procedures and definitions to construct the moments in data and in model simulations.

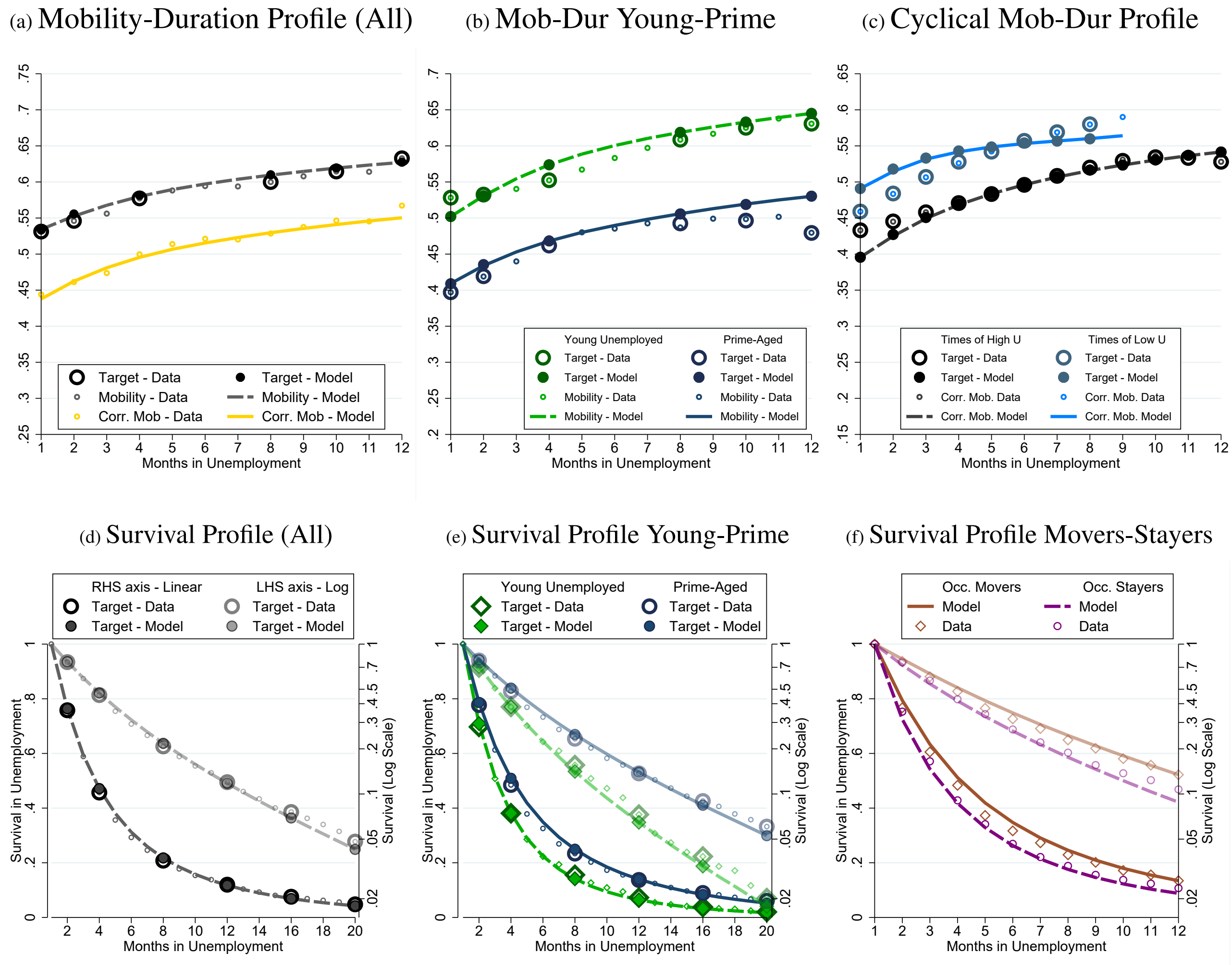
We now present some heuristic identification arguments that justify our choice of moments, keeping in mind that all parameters need to be estimated jointly. For ease of exposition we first discuss those moments that inform parameters that are shared by all occupations and then those moments that inform occupation-specific parameters. For the former we use occupational mobility statistics based on the major occupation classification, while for the latter we use the task-based classification.

Economy-wide moments A worker's attachment to his pre-separation occupation during an unemployment spell depends on the properties of the z-productivity process, the human capital process and the reallocation cost c. The aggregate and age-group mobility-duration profiles depicted in Figures 8a and 8b (see also Section 2) play an important role in informing these parameters. In particular, the aggregate mobility-duration profile contains information about c and ρ_z. As shown in Lemma 1 (Appendix B.1) changes in the overall level of mobility lead to opposite changes in c. The slope of the mobility-duration profile informs ρ_z primarily through the time it takes unemployed workers to start searching across occupations.[21] A lower ρ_z (keeping constant $F(z)$) increases the relative number of unemployed workers deciding to search across occupations at shorter durations, decreases the slope of the model's aggregate mobility-duration profile. Lemma 1, however, also implies that a lower ρ_z reduces overall mobility (ceteris paribus), creating a tension between c and ρ_z such that an increase in ρ_z must go together with an increase in c to fit the observed mobility-duration profile. To help identify σ_z we use instead the mobility-duration profiles of young and prime-aged workers. For given values of x_h, a larger value of σ_z leads to a smaller importance of human capital differences relative to z-productivities differences in workers' output. This brings the simulated occupational mobility patterns across age groups closer together, creating a negative relationship between σ_z and the difference between the mobility-duration profiles of young and prime-aged workers.

To inform the human capital parameters x_2 and x_3 we use the overall level of occupational mobility among young and prime-aged workers (see Lemma 2, Appendix B.1) as well as the observed five and ten-year returns to occupational experience. Since it is difficult to estimate returns to occupational

[21] Other factors that might allow the calibrated model to generate the observed mobility-duration profile do not appear important in our estimation. In particular, the large extent of occupational mobility at short unemployment durations implies that the time it takes a typical worker to decide to search in a given occupation is small and hence does not drive the observed unemployment duration differences between occupational movers and stayers. Further, since the structure of the model implies that exogenous separated workers and occupational movers have very similar realised z-distributions, composition effects in post-reallocation outcomes do not play an important role. Finally, the changes in the mean-reversion of the z-productivity process brought about by changes in ρ_z seem to only play a minor role in shaping the mobility-duration profile.

Figure 8: Targeted Moments. Data and Model Comparison



experience accurately with the SIPP due to the relative short nature of its panels, we use the OLS estimates for 1-digit occupations reported in Kambourov and Manovskii (2009b) from the PSID and estimate the same OLS regression in simulated data. We use the OLS estimates because occupation selection occurs both in the model and in the data, where selection arises as measured returns are a result of two opposing forces: human capital acquisition and z-productivity mean reversion.

A worker's attachment to employment depends on the size of search frictions. A higher value of k leads to stronger search frictions through its effect on firm entry and labor market tightness. This pushes down the z^s cutoffs and reduces the extent of endogenous separations. Thus, to recover k, we can use information on the importance of endogenous separations. Since a persistent z process implies endogenous separations beget future endogenous separations, to inform the extent of endogenous separations and hence k we use as targets the proportion of separations within a year of leaving unemployment relative to the overall yearly separation rate ("Rel. separation rate recent hire/all"), and the concentration of unemployment spells over a SIPP panel window among the subset of workers who start employed at the beginning of the panel ("Prob (unemp. within 3yr for empl.)"). The probability that an occupational stayer becomes an occupational mover in the next unemployment spell ("Repeat mobility") further informs endogenous separations and how these relate to occupational

Table 3: Targeted Moments. Data and Model Comparison

Panel A: Economy-wide moments

Moment	Model	Data	Moment	Model	Data
Agg. output per worker mean	0.999	1.000	Rel. separation rate young/prime-aged	1.999	2.044
Agg. output per worker persistence, ρ_{outpw}	0.764	0.753	Rel. separation rate recent hire/all	5.180	4.945
Agg. output per worker st. dev., σ_{outpw}	0.009	0.009	Prob (unemp. within 3 yr for empl.)	0.151	0.124
Mean unemployment	0.036	0.036	Empirical elasticity matching function	0.526	0.500
Task-based gross occ. mobility rate	0.280	0.288	5-year OLS return to occ. tenure	0.143	0.154
Repeat mobility: occ. stay after stay	0.600	0.649	10-year OLS return to occ. tenure	0.219	0.232
Occ. mobility young/prime-aged	1.167	1.163	Average u. duration movers/stayers	1.181	1.140
Occ. mobility-duration profiles:	Fig 7a,b,c		U. survival profiles	Fig 7d,e	

Panel B: Occupation-Specific Moments, Long-run

	Proportion empl. size o_{2014}		Net mobility *Mean*		Transition Matrix Model				Transition Matrix Data			
	Model	Data	Model	Data	NRC	RC	NRM	RM	NRC	RC	NRM	RM
NRC	0.337	0.329	0.008	0.006	0.763	0.164	0.055	0.018	0.722	0.167	0.084	0.028
RC	0.246	0.258	0.006	0.001	0.129	0.681	0.144	0.047	0.078	0.681	0.168	0.066
NRM	0.260	0.260	-0.027	-0.021	0.034	0.065	0.760	0.141	0.020	0.115	0.710	0.155
RM	0.157	0.154	0.011	0.015	0.037	0.069	0.247	0.647	0.013	0.066	0.188	0.733

Panel C: Occupation-Specific Moments, Cyclical

	Net mobility *Recessions*		Net mobility *Expansions*		Net mobility *Rec-Exp*		$\Delta_{exp-rec}$ (inflow o/all flows)		$\varepsilon_{UD_o,u}/\varepsilon_{UD_{avg},u}$	
	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data
NRC	-0.012	-0.011	-0.002	-0.003	-0.010	-0.008	-0.003	-0.010	0.996	1.096
RC	-0.009	-0.005	-0.005	-0.001	-0.004	-0.004	0.006	0.003	1.054	1.027
NRM	0.034	0.033	0.017	0.011	0.017	0.022	-0.066	-0.054	0.874	0.761
RM	-0.017	-0.017	-0.006	-0.008	-0.011	-0.009	0.027	0.061	1.081	1.122

mobility.

The elasticity of the matching function on each labor market, η, is obtained by estimating through OLS a log-linear relation between the aggregate job finding rate (the proportion of all unemployed who have a job next month) and aggregate labor market tightness (all vacancies over all unemployed) across quarters, in simulated data. The estimated elasticity $\hat{\eta}$ is targeted to be 0.5 (see Petrongolo and Pissarides, 2001) and allows us to indirectly infer η.

We use the unemployment survival function depicted in Figure 8d to additionally inform us about all the above parameters. The extent of duration dependence is linked to the properties of the z process and the importance of search frictions through their effect on the extent of true duration dependence and dynamic selection in our model, where the latter is driven by worker heterogeneity in x and z at the moment of separation. We use the cumulative survival rates at intervals of 4 months to reduce the seam bias found in the SIPP. To further emphasize the difference between the average unemployment durations of occupational movers and stayers, we also target its ratio.

Given these job-finding moments, the overall separation rate follows from targeting the unemployment rate. As we focus on those who held a job previously, we use the most direct counterpart and construct the unemployment rate only for those who were employed before and satisfied our definition of unemployment (see Section 2). Note that this unemployment rate (3.6%) is lower than the BLS unemployment rate, but we find it responsible for more than 0.75 for every one percentage point change in the BLS unemployment rate (see Appendix C.1 and Supplementary Appendix G for details), consistent with the results of Hornstein (2013), Fujita and Moscarini (2017) and Ahn and Hamilton (2018).

The ratio of separation rates between young and prime-aged workers ("Rel. separation rate young/prime-aged") as well as their survival functions in Figure 8e inform δ_L, δ_H and b. A relatively higher δ_L shifts the realised z-distribution of newly separated inexperienced workers away from their z^s cutoff towards higher z levels and hence affect their extent of duration dependence in unemployment, especially at shorter durations. The extent of separations for young and prime-aged workers also informs us about b through the positions of the z^s cutoffs of low and high human capital workers relative to average of these workers' productivities.

Differently from all the previous structural parameters, calibrations with or without occupational human capital depreciation yield very similar long-run moments (see Appendix C.2.2). This occurs because the gradual loss of occupational attachment with unemployment duration underlying the observed mobility-duration profile can be generated by human capital depreciation or the z process. To differentiate these two forces we instead use the cyclical shift of the mobility-duration profile. During recessions longer unemployment spells imply that expected human capital depreciation is higher, making employed workers more attached to their jobs and unemployed workers less attached to their occupations. At the same time low aggregate productivity interacted with the z-productivities typically makes employed workers less attached to their jobs and unemployed workers more attached to their occupations. To inform this tension and recover γ_d we target the mobility-duration profile in recessions and expansions as depicted in Figure 8c (see also Section 2).

The persistence and standard deviation of the aggregate productivity process, ρ_A and σ_A, are informed by the corresponding parameters of the series of output per worker ($outpw$) obtained from the BLS, ρ_{outpw} and σ_{outpw}, and measured quarterly for the period 1983-2014.[22]

Occupation-specific moments Variation over the business cycle can also naturally inform the cyclical sensitivity of occupation-wide productivities and the degree of directness in workers' search across occupations. In particular, to inform ϵ_o we target the levels of net mobility each task-based category $o \in \{NRC, RC, NRM, RM\}$ exhibits in recession and expansions ("Net mobility o, *Recessions* and Net mobility o, *Expansions*") as depicted in Figure 6 (Section 2) as well as their implied difference ("Net mobility o, *Rec-Exp*"). In addition, for each o we regress the completed (log) unemployment durations of those workers whose pre-separation task-based category was o on the (log) unemployment rate and a time trend, and target the ratio between the estimated unemployment duration elasticity and the average elasticity across task-based categories, $\varepsilon_{UD_o,u}/\varepsilon_{UD_{avg},u}$ (see Appendix C.1 for details).

[22] We cannot set these parameters directly because the composition of the economy changes with the cycle due to workers' endogenous separation and reallocation decisions. We measure output in the model and data on a quarterly basis (aggregating the underlying weekly process in the model). For the data, we HP-filtered the series of (log) output per worker for the period 1970 to 2016. Then, we use the persistence and the variance parameters of this series calculated over the period 1983-2014, which is the period that the SIPP and the BLS series overlap.

The advantage of this approach is that it allows us to leave untargeted the cyclicality of aggregate unemployment, which we separately evaluate in Section 5. To inform the values of $\overline{p}_o$ we target the average net mobility level of each o across the cycle ("Net mobility o, *Mean*").

To recover ν we exploit the observed differences in the cyclicality of inflows across task-based categories. As ν increases, workers should be more sensitive (ceteris paribus) to cyclical differences in p_o when choosing occupations, making the inflows to occupations with the higher p_o respond stronger. To capture how cyclically sensitive are the inflows we compute, separately for expansions and recessions, the ratio of inflows into task-based category o over the sum of all flows. For each o we target the difference between the expansion and recession ratios, $\Delta_{exp-rec}$ (inflow $o/$all flows). The values of $\overline{\alpha}_{o,\tilde{o}}$ are informed by targeting the observed task-based occupation transition matrix.

To recover the set of ψ_o we use the employment-size distribution of task-based categories observed in 2014, the end of our sample period, "Prop (empl. size o_{2014})". We also target the average gross mobility rate across task-based categories ("Task-based gross occ. mobility rate") so that the model remains consistent with gross mobility at this level of aggregation.

Table 4: Calibrated Parameters

Agg. prod. and search frictions		ρ_A	σ_A	b	k	η	
		0.9985	0.0020	0.830	124.83	0.239	
Occ. human capital process		x_2	x_3	γ_d	δ_L	δ_H	
		1.171	1.458	0.0032	0.0035	0.0002	
Occupational mobility		c	ρ_z	σ_z	z_{norm}	ν	
		7.603	0.9983	0.0072	0.354	0.04	
Occupation-specific	$\overline{p}_o$	ϵ_o	ψ_o	$\overline{\alpha}_{o,NRC}$	$\overline{\alpha}_{o,RC}$	$\overline{\alpha}_{o,NRM}$	$\overline{\alpha}_{o,RM}$
Non-routine Cognitive	1.019	1.082	0.620	0.436	0.560	0.004	0.000
Routine Cognitive	0.988	1.120	0.145	0.407	0.383	0.210	0.000
Non-routine Manual	1.000	0.532	0.087	0.000	0.093	0.384	0.524
Routine Manual	0.988	1.283	0.147	0.000	0.140	0.767	0.094

Parameters Table 4 reports the resulting parameter values implied by the calibration. The estimated value of b represents about 80% of total average output, y, not too far off from Hall and Milgrom's (2008) estimate, though we use different information. Vacancy cost k translates to a cost of about 30% of weekly output to fill a job. The elasticity of the matching function within each labor market $\eta = 0.24$ which is about half of $\widehat{\eta} = 0.5$, when aggregating across all labor markets.[23]

The actual returns to occupational experience x_2 and x_3 are higher than the OLS returns, because occupational entrants select better z-productivities that typically mean-revert over time, dampening the average evolution of *composite* xz-productivity. The parameter γ_d implies that a year in unemployment costs an experienced worker in expectation about 5% of his productivity. The estimated

[23] The difference between η and $\widehat{\eta}$ is mainly due to the effect of aggregation across submarkets that exhibit rest unemployment (below we show that rest unemployment episodes occur in our calibration). Workers in episodes of rest unemployed entail no vacancies, have zero job finding rates, do not congest matching in other submarket, but are included in the aggregate number of unemployed. For the latter reason rest unemployed workers show up in the empirical matching function and drive a wedge between η and 0.5. Moreover, each submarket exhibits a concave matching function and hence aggregation over these concave functions also imply that the calibrated value of η will further diverge from 0.5.

values of δ_L and δ_H imply that exogenous separations are much more prevalent for low rather than high human capital workers, leading to a larger importance of endogenous separations among the latter, as implied by the prime-aged survival and mobility-duration profiles. The estimated value of c and the sampling process imply that upon starting a job in a new occupation, a worker has paid on average a reallocation cost of 15.18 weeks (or about 3.5 months) of output. This suggests that reallocation frictions are important and add to the significant lose in occupational human capital when changing occupation.[24]

The process driving workers' idiosyncratic productivities within an occupation has a broadly similar persistence (at a weekly basis) as the aggregate shock process driving the business cycle. However, its larger variance implies there is much more dispersion across workers' z-productivities than there is across values of A. We also find that workers' idiosyncratic productivities are much more dispersed than occupation-wide productivities. For example, the max-min ratio of p_o is 1.13 (1.09) at the highest (lowest) value of A, where the RM task-based category is the most responsive to aggregate shocks and NRM the least. In contrast, the max-min ratio among z-productivities is 2.20. To gauge whether the dispersion across z-productivities is reasonable we calculate the implied amount of frictional wage dispersion using Hornstein et al. (2012) Mm ratio. These authors find an Mm between 1.46 and 1.90 using the PSID, while the estimated z-dispersion yields 1.40.

The estimated value of ν implies that the ability of workers to access job opportunities in other task-based categories plays an important role in shaping the direction of their search. The estimated values of $\overline{\alpha}_{o,\tilde{o}}$ imply that on average workers in NRC have a low probability of drawing a new z from manual occupations and vice versa; while workers in NRM and RM categories mostly draw a new z from these same two categories, although drawing from RC is not uncommon. As discussed in the next subsection, however, workers do react to cyclical occupation-wide productivity differences when deciding which task-based category to search in. This is evidenced by the ability of the model to reproduce the observed cyclical changes in the net mobility patterns presented in Section 2 and Table 3 (see also Appendix C.1). Further, Table 7 in Section 5 shows an important role of occupational mobility through unemployment in changing the relative sizes of NRM and RM occupations. In contrast, the high value of ψ_{NRC} captures that the NRC category did not increase its size between 1984 and 2014 because of inflows through unemployment, but rather because of a significant proportion of labor market entrants taking up jobs there (see Table 7 below).

4.2 Fit of the Model

Figure 8 and Table 3 show that the model fits the data very well, given the extent of over-identification. Crucially the calibrated model is able to generate procyclical gross occupational mobility and coun-

[24]The average reallocation cost is computed as the product of c and the number of times workers sample a new occupation, which is 1.996 times. The value of c reported in Table 4 is consistent with the large proportion of unemployed workers who change occupation. Given that in the data occupational changes are typically accompanied by changes in industries (based in our own calculations) and, to a lesser extent, by geographical location (see Papageorgiou, 2018), the estimated value of c could also be capturing the moving costs associated with these changes. Indeed, Alvarez and Shimer (2011) find also large reallocation costs across industries, while Kennan and Walker (2011) and Papageorgiou (2018) find large reallocation costs across geographical locations.

tercyclical net mobility as in the data together with the long-run moments that characterise net and gross mobility documented in Section 2.

In particular, it captures that the probability of gross occupational mobility falls in recessions at both short and long unemployment durations, and that a higher proportion of those unemployed who nevertheless changed occupations in recessions contribute further to the decrease in the size of the RM, NRC and RC categories and to the increase in the size of the NRM category. Across all five mobility-duration profiles, the model reproduces the differing high levels of occupational mobility and the shared moderate increase with duration. Human capital accumulation (and depreciation) allows the model to reproduce the largely parallel downward shift of the mobility-duration profile across age-groups. In Appendix C.1 we show that the calibrated model also remains fully consistent with the much larger contribution of excess mobility relative to net mobility in accounting for the mobility-duration profile at all durations, as depicted in Figure 3b. The model also captures well the probability that a worker who did not change occupations after an unemployment spell, did not change occupation after a subsequent unemployment spell. It is also inline with the (untargeted) probability that a worker who changed occupation after an unemployment spell, changed occupation after a subsequent unemployment spell. This probability is 0.54 in the model and 0.56 in the data. The model also remains consistent with the long-run changes in the employment size distribution observed across task-based categories.

The calibration further generates aggregate and age-specific (young and prime-aged) survival functions consistent with the data, matching the observed duration dependence of unemployment. In Appendix C.1 we show that this leads the model to match well the (untargeted) incomplete unemployment duration distribution among all workers and separately by age groups, in particular the empirical amount of long-term unemployment that occurs in the face of high occupational mobility. We also observe that both in the model and the data duration dependence is different across (ex post) occupational stayers and movers and across age groups, where duration dependence is stronger among occupational stayers relative to movers and among young relative to prime-aged workers. Young occupational stayers have especially high job finding at low durations, which decrease faster as duration increases. In Appendix C.1 we show that the calibrated model reproduces well the (untargeted) hazard functions associated with the above survival functions.

The calibration thus shows that the z-productivity process allows the model to reproduce very well a wide-range of long-run gross occupational mobility and unemployment duration moments. Interacted with aggregated productivity it generates the observed cyclical shifts in the mobility-duration profile. Differences in occupation-wide productivities loadings, ϵ_o, and the responsiveness of the direction of search across occupations, ν, further allow the model to be consistent with the cyclical changes in net mobility for each task-based category as well as their relative differences in cyclical unemployment duration responses (see Appendix C.1).

Finally, the z-productivity process and the implied endogenous separations also allow the model to capture that unemployment risk is concentrated among a subset of workers in the data. The persistence in the z process means that workers hired from unemployment are often in z-productivities close to the separation cutoffs, leading to a yearly separation rate that is about 5 times higher than for the average employed worker both in the data and model. In the data and model we also find that the probability an employed worker at the beginning of a SIPP panel becomes unemployed in the next 3 years is lower than the one that would be generated if we imposed the SIPP's monthly job separation rate across all workers. Although not targeted, the model captures that more initially employed young workers would become unemployment in the next 3 years than prime-aged workers (7 p.p in the data vs. 8 p.p in the model). This occurs because of the larger importance of endogenous separations among high human capital, prime-aged workers.

5 Cyclical Unemployment Outcomes

We now turn to investigate the cyclical patterns of aggregate unemployment and its duration distribution generated by the model, noting that these were not targeted in our estimation procedure. Our aim is to investigate the importance of excess and net occupational mobility in generating these patterns. We first present the implications of the full model as estimated above. We then discuss the implications of a re-estimated version of the model where we shut down endogenous net mobility by not allowing workers to decide in which occupations to search in and dropping occupation-wide productivity differences.[25] With a slight abuse of terminology, we label this version "excess mobility model" as unemployed workers' occupational mobility decisions are based solely on the changing nature of their z-productivities and their interaction with A and x_h. In Appendix C.2 we present the estimation results of the excess mobility model.

Table 5: Logged and HP-filtered Business Cycle Statistics. Data (1983-2014) and Model

	Volatility and Persistence							Correlations with u and outpw							
	u	v	θ	s	f	$outpw$	$occm$		u	v	θ	s	f	$outpw$	$occm$
Data															
σ	0.14	0.11	0.25	0.10	0.09	0.01	0.03	u	1.00	-0.92	-0.98	0.80	-0.82	-0.47	-0.52
ρ_{t-1}	0.98	0.99	0.99	0.94	0.91	0.93	0.91	$outpw$		0.56	0.51	-0.39	0.27	1.00	0.38
Full Model															
σ	0.14	0.05	0.17	0.07	0.10	0.01	0.04	u	1.00	-0.61	-0.96	0.79	-0.88	-0.94	-0.82
ρ_{t-1}	0.93	0.90	0.92	0.87	0.92	0.88	0.93	$outpw$		0.76	0.96	-0.90	0.93	1.00	0.83
Exc. Mob. Model															
σ	0.14	0.05	0.18	0.07	0.10	0.01	0.04	u	1.00	-0.63	-0.97	0.78	-0.88	-0.94	-0.80
ρ_{t-1}	0.95	0.89	0.94	0.88	0.93	0.94	0.90	$outpw$		0.77	0.96	-0.87	0.93	1.00	0.83

Note: The excess mobility model considers only occ. mobility decisions based on the z-productivity process. Each model's aggregate time series arise from the distributions of employed and unemployed workers across all labor markets, combined with agents' decisions. Times series are centered 5Q-MA series of quarterly data (both model and data), to smooth out the discreteness in the relatively flat cutoffs (relative to the grid) discussed further in the computational appendix. The cyclical components of the (log) of these time series are obtained by using an HP filter with parameter 1600. See Appendix C.1 for further details and results without the 5Q-MA smoothing.

Aggregate unemployment Table 5 shows the cyclical properties of the aggregate unemployment, vacancy, job finding and separation and gross occupational mobility rates, computed from the data

[25] In this version the observed net mobility patterns can be imposed exogenously to keep the model's gross occupational mobility patterns consistent with the evidence presented in Section 2 and the Supplementary Appendix. An alternative exercise would be to maintain productivity differences across occupations but not allow workers to chose in which occupations to search on. Given that the estimated dispersion of z-productivities is much larger than that of p_o productivities, this exercise would not generate meaningfully different results. A second alternative could be to re-estimate a version of the model where we shut down the z-productivity process, making workers decide whether to change occupations based only on p_o productivities differences. It is clear, however, that this version of the model will not be able to reproduce many of the occupational mobility patterns documented in Section 2.

Panel A: Cyclicality of Duration Distribution						
Unemp.	Elasticity wrt u			Semi-elasticity wrt u		
Duration	Full Model	Excess Model	Data	Full Model	Excess Model	Data
$1-2$m	-0.435	-0.447	-0.464	-0.168	-0.165	-0.169
$1-4$m	-0.316	-0.329	-0.363	-0.178	-0.179	-0.186
$5-8$m	0.388	0.350	0.320	0.074	0.070	0.071
$9-12$m	1.083	1.033	0.864	0.061	0.060	0.072
> 13m	1.787	1.513	1.375	0.047	0.048	0.044

Panel B: Semi-Elasticity Duration wrt u by Occupational Mobility						
Unemp.	HP-filtered			Log u linearly detrended		
Duration by Mob.	Full Model	Excess Model	Data	Full Model	Excess Model	Data
Movers	2.9	2.9	3.2	2.4	2.3	2.0
Stayers	1.5	1.4	2.5	1.2	1.2	1.6

Note: The elasticities are constructed using the cyclical component of the series of the shares of unemployed workers by durations, the aggregate unemployment rate.

Table 6: Cyclical duration distribution

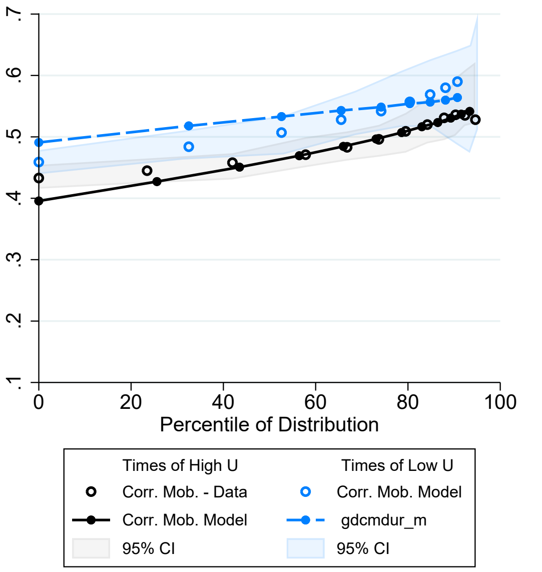


Figure 9: Cyclical Shift of Distribution

and the simulations.[26] It shows that the full model is able to generate a countercyclical unemployment rate, together with a countercyclical job separation rate, procyclical job finding and gross occupational mobility rates and a strongly negatively-sloped Beveridge curve as in the data. The latter stands in contrast with the canonical DMP model, where it is known that endogenous separations hamper this model from achieving a Beveridge curve consistent with the data. Table 5 also shows that the cyclical volatilities and persistence of the aggregate unemployment, job finding, separation and gross occupational mobility rates are very close to the data.

Note that this aggregate behavior is not driven by a higher cyclicality of young workers' unemployment rate. In Appendix C.1 we show that the responsiveness of the unemployment rate to aggregate output per worker is slightly stronger for prime-aged workers than for young workers, leading to a countercyclical ratio of unemployment rates between young and prime-aged workers. Therefore, in the model the pool of unemployment shifts towards high human-capital, prime-aged workers during recessions, a feature noted by Mueller (2017). In the model this occurs mostly due to the larger increase in endogenous job separations among prime-aged relative to young workers.

Unemployment duration distribution Panel A in Table 6 evaluates the ability of the model to reproduce the shifts in the incomplete unemployment duration distribution with respect to changes in the unemployment rate. It shows that the shares of unemployed workers by duration exhibit a very similar degree of responsiveness with cyclical unemployment as in the data. Crucially the elasticity measure shows that the model creates a strong response in the shares of unemployment at long durations. When using the semi-elasticity measure the model generates a nearly perfect fit. Thus, in our model as in the data cyclical changes in the aggregate unemployment rate are driven by particularly strong cyclical changes in long-term unemployment.

[26] Both in the model and data the unemployment, job finding and separation rates are computed based on the same unemployment definition used in the previous sections, while the cyclical properties of the occupational mobility rate are computed using major occupational groups and after applying the $\mathbf{\Gamma}$-correction matrix. In Appendix C.1 we provide the full set of correlations.

An important force behind the increase in long-term unemployment during recessions is the larger increase in the unemployment duration of occupational movers relative to stayers. Panel B in Table 6 shows the cyclical responses of the average unemployment duration of movers and stayers using different measures. Along all of these measures the model reproduces that the average unemployment duration of occupational movers increases more than the average unemployment duration of stayers, with an increase that is consistent with the data. Stayers' durations respond somewhat less relative to the data, between 60% (relative to the log HP-filtered unemployment measure) and 80% (relative to the linearly detrended unemployment measure). Relative to the lack of amplification in conventional DMP models, this still constitutes a large response. As in the data, the lengthening of movers' unemployment duration contributes meaningfully to the increase in long-term unemployment during recessions.

Figure 9 (see also Figure 5b in Section 2) shows how the untargeted shift in unemployment durations combines with the targeted shift of the mobility-duration profile. At any percentile of the unemployment duration distribution, the model generates a drop in occupational mobility in recessions. By comparing the observations' x-coordinates, this figure also illustrates that the cyclical shift of the model's duration distribution follows the data.

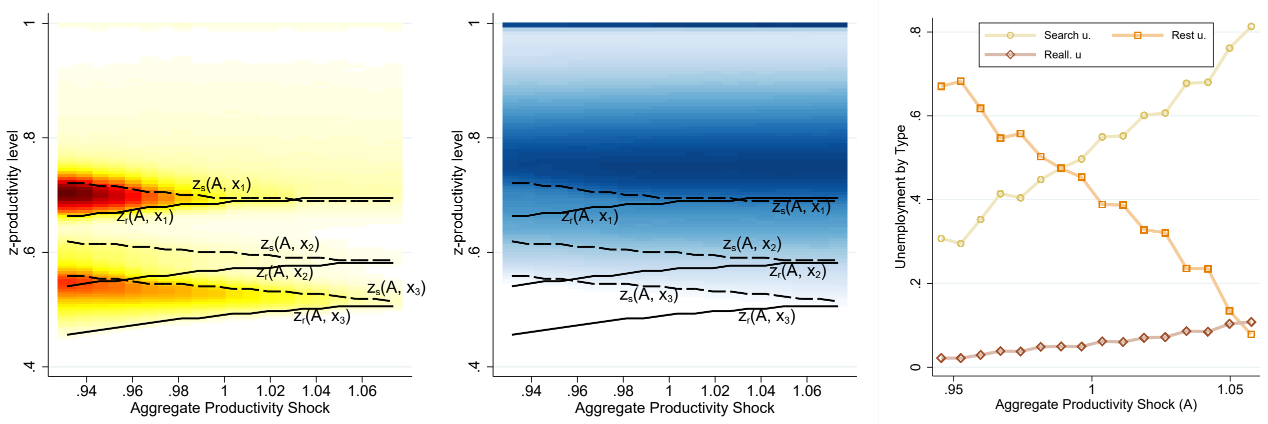
Excess vs. net mobility A key insight from Tables 5 and 6 is that the aforementioned cyclical patterns are nearly identical to the ones generated by the excess mobility model. In Appendix C.2 we show that this model also fits very well the economy-wide targets discussed in Section 4 and the estimated values of $[c, \rho_z, \sigma_z, \underline{z}_{norm}, x_2, x_3, \gamma_d, \delta_L, \delta_H, k, b, \eta, \rho_A, \sigma_A]$ are nearly identical to the estimated in the full model. This comparison demonstrates that allowing workers to chose in which occupations to search in due to occupation-wide productivity differences is not the reason why the model is able to replicate the cyclical patterns of aggregate unemployment and its duration distribution. Instead, the excess mobility calibration highlights the importance of the worker-occupation idiosyncratic productivity process and its interaction with aggregate productivity in generating these cyclical patterns.

The excess mobility model and the full model calibrations are successful in these dimensions because they yield similar implications for search, rest and reallocation unemployment during workers' unemployment spells. For ease of exposition, in Section 5.1 we first demonstrate this claim using the excess mobility model calibration. In Section 5.2 we show that the same forces occur within each task-based category in the full model, although modulated by differences in the level and cyclical responsiveness of p_o across occupations. Below we also demonstrate that it is crucial for the empirical success of the model to allow for occupational mobility and not only for search and rest unemployment episodes.

5.1 The Co-existence of Search and Rest Unemployment

As argued in Section 3.3, the relative position and slopes of z^s and z^r are key determinants of the long-run and cyclical implications of our model. We now discuss these in the context of the calibrations.

Relative position of z^s and z^r Figure 10a depicts the cutoff functions generated by the excess model calibration as a function A given x_h, where all occupations share the same cutoff functions.



(a) Unemployment Distribution (b) Employment Distribution (c) Unemployment Decomposition

Figure 10: Cutoffs, Unemployment Distribution and Decomposition

It shows that $z^s \geq z^r$ for nearly all A and $h = 1, 2, 3$. The exception being $z^s(A; x_1) < z^r(A; x_1)$ for the highest values of A. This implies that periods of search, rest and reallocation unemployment can occur within the same unemployment spell as A and z evolve. Further z^s and z^r decrease with x_h such that, as predicted by our theory, workers with higher human capital are less likely to change occupations relative to those with lower human capital. As $z^s(., x_3) < z^s(., x_1)$ the average level of separations is also lower for high human capital workers (noting that δ_L and δ_H also contribute to this difference). Once separated high human capital workers spent on average a longer time in unemployment due to the larger distance between their z^s and z^r cutoffs.

Given the values of x_h, our theory shows that c, b and ρ_z determine the distance between z^s and z^r, and therefore the presence of episodes of rest unemployment. These parameters are informed by the mobility-duration profiles, the unemployment survival functions, repeat mobility patterns and the incidence of unemployment among employed workers. To illustrate why parameter values that lead to $z^s \geq z^r$ allow the model to match the mobility-duration profile and survival functions, consider a set of workers with the same x_h who just endogenously separated. Given $z^s \geq z^r$ and a z process close to a random walk (as in the calibration), these workers will be initially close to z^s. A small positive shock would then suffice to move them above z^s, while only large negative shocks would take them below z^r. Hence at short durations these workers face relatively high job finding rates and, if re-employed, they will be most likely occupational stayers. Those who stayed unemployed for longer would have then experienced further negative z shocks and would face a higher probability of crossing z^r. At the same time, even for the long-term unemployed the probability of exiting rest unemployment by crossing z^s remains relevant as they face a reasonably high probability of re-employment within their occupation. As a result, the likelihood of an occupational move increases with unemployment duration, while the job finding rate decreases with unemployment duration, consistent with the observed mobility-duration profiles and the unemployment survivor functions.

With $z^s \geq z^r$ and prominent endogenous job separations (as in the calibration), employed workers close to z^s would face a high risk of separating. Upon becoming re-employed in the same occupation, these workers would still be close to z^s, facing once again a high job separation probability. Occupational movers would be distributed over the entire range of z-productivities above the z^s cutoff in the new occupation. Among this set of workers, those close to the z^s cutoff face once again a high risk of unemployment and subsequent repeat mobility patterns, inline with the data.

Slope of z^s and z^r Figure 10a shows that $\partial z^s/\partial A < 0$ and $\partial z^r/\partial A > 0$ for each x_h. This property implies that recessions are periods in which the two cutoffs become further apart, increasing the scope for episodes of rest unemployment; while expansions are periods in which these cutoffs become closer, increasing the chances for episodes of search unemployment. Figure 10c illustrate this last feature by showing the proportion of workers facing an episode of search, rest or reallocation unemployment for a given value of A. Although both rest and search unemployment are countercyclical, search unemployment episodes are relatively more common when the economy moves from mild recessions up to strong expansions. It is only as recessions get stronger that rest unemployment episodes become more common. These dynamics highlight that the intervals $[z^r(A; x_h), z^s(A; x_h)]$ create an "area of inaction" for each x_h that widens in recession, trapping workers in rest unemployment episodes who consequently leave the areas of inaction at a lower rate. As conditions improve these areas narrow considerably such that rest unemployed workers are now much more likely to get a z shock that takes them below (or above) their z^r (z^s) cutoff. Indeed, in recessions that involve a 5% reduction in A relative to the mean, workers still face an average probability of about 25% of transitioning out of rest unemployment within a month; and this probability sharply increases with aggregate productivity.

The cyclicality of aggregate unemployment in the calibration is then determined by mainly four factors: (i) the values of ρ_z and σ_z; (ii) the steepness of z^s and z^r; (iii) the distribution of employed and unemployed workers over z-productivities; and (iv) human capital depreciation. Factor (i) can be gauged from Table 4, while Figures 10a and 10b present a heat map depicting (ii) and (iii) for a given value of A. We discuss factor (iv) below.

The negative slope of the z^s cutoff together with a large mass of workers right above it imply that a decrease in A leads to a large *increase* in the *inflow* of workers into rest unemployment episodes. The same decrease in A also leads to a large *decrease* in the *outflow* from rest unemployment due to the widening of the areas of inaction, $[z^r(A; x_h), z^s(A; x_h)]$, and the increase in the density of unemployed workers at any given z within these areas. The latter captures that in a recession, the mass of workers in a rest unemployment episode at a given z includes those who would have changed occupations in normal times, but now prefer to wait for conditions to improve. As aggregate conditions improve workers are more likely to experience episodes of search unemployment by crossing both $z^r(A; x_h)$ and $z^s(A; x_h)$. Given the strong responses in rest and search unemployment, u also becomes highly responsive to A. Episodes of reallocation unemployment, however, make a small contribution to the cyclicality of u because they only capture the time spent transiting between occupations, which is about 2 weeks on average, after which workers continue their jobless spell in episodes of rest or search before finding a job in a new occupation. In Appendix C.2 we show that these patterns occur across low and high human capital levels, explaining why we obtain unemployment, job finding and separations rates across age groups with similar cyclical responses.

The cyclical change in the areas of inaction also enable the model to obtain a strong Beveridge curve. With downward sloping and relatively steep z^s cutoffs, a decrease in A yields a decrease in the amount of vacancies being creating above these cutoffs as match surpluses shrink. The same decrease in A implies that more z-productivities lie below z^s and no vacancies are created in that region leading to a further decrease in v. As A increases, the opposite happens. Procyclical occupational mobility also generates procyclical vacancy creation, as workers who search in different occupation end up with $z > z^s$ and firms respond by creating more vacancies. Together, these forces imply that vacancies become responsive to A and exhibit a high negative correlation with u.

The fanning out of the area of inaction during recessions also drives the cyclical behavior of the unemployment duration distribution. As A decreases, the distance for the typical long-term unemployed worker to both cutoffs is substantially larger. At low values of A, long-term unemployed workers will require a sequence of more and larger good shocks to their z-productivities before becoming search unemployed in their pre-separation occupation. They would also require a sequence of more and larger bad shocks to their z-productivities before deciding to change occupations. In contrast, workers who have just endogenously separated tend to be close to their respective z^s cutoffs at all values of A. Since in recessions z^r is further away for these workers, z^s is the cutoff that weighs most on their future outcomes. This implies that for workers who have just endogenously separated, the distance to the nearest cutoff is not as responsive to A as for the long-term unemployed. Hence, we observe that the outflow rate of long-term unemployed workers responds more to changes in aggregate conditions relative to the outflow rate of shorter-term unemployed workers. This mechanism then translates into a stronger increase in the share of long-term unemployed in recessions as shown in Table 6, stronger than the one predicted based on the decline of f alone. The same mechanism also implies that at low values of A the time spent in rest unemployment increases more for (ex-post) occupational movers than for occupational stayers, rationalizing the stronger increase in average unemployment duration among occupational movers relative to stayers during recessions.

Human capital depreciation is important in determining the cyclical behavior of aggregate unemployment and its duration distribution because it affects the cyclical changes in the areas of inaction. Workers with a z-productivity much lower than z^s take into account that even with a sequence of positive z realizations they might experience depreciation and reallocate anyway, decreasing the option value of waiting in their occupations and flattening the z^r cutoff. At separation a similar argument operates: increases in $z^s - z^r$ during recessions implies that depreciation more often leads to a reallocation than otherwise, increasing the option value of staying employed and flattening the z^s cutoff. In Appendix C.2.2 we show that this mechanism leads to a stronger amplification of rest unemployment when we do not allow for human capital depreciation. As a consequence, such a version of the model generates too large a volatility of the aggregate unemployment rate as well as too little occupational mobility during recessions.

The importance of occupational mobility To gauge the importance of the reallocation cutoff on the cyclical performance of our model we re-estimate the model not allowing workers to change occupations. We use all the same moments outlined before except those pertaining to occupational mobility. In Appendix C.3 we show that the calibrated one-sector model with no occupational mobility can do well in fitting most of the targeted long-run moments, particularly the unemployment survival functions for all workers and by age groups. However, the aggregate unemployment, vacancy, job finding and separation rates now exhibit below half the cyclical volatility observed in their data counterparts, 0.04, 0.02, 0.03 and 0.03, respectively, and the correlation between unemployment and vacancies drops to -0.32 (see Table 5 for a comparison).

The main reason why this version generates such a low cyclical response is that the new area of inaction is defined by the set of z-productivities that lie below the z^s cutoff down to the lowest value of the z-productivities. Crucially, any cyclical changes in the size of this area now solely depend on the responsiveness of the z^s cutoff relative to the workers' z distribution. Although $\partial z^s/\partial A < 0$ and hence this calibration generates countercyclical separations, it cannot resolve a key trade-off: in the absence of the z^r cutoff we now estimate a less persistent and much more volatile z process which creates enough heterogeneity in unemployment durations that allows it to match the empirical unemployment survival functions. However, the estimated properties of the z process also increase the heterogeneity in z-productivities relative to the cyclical range of A. This dampens the model's cyclical performance as it implies less responsive z^s cutoffs, weakening the cyclical responses of job separations and the rate at which rest unemployed workers leave the area of inaction.

In Appendix C.3 we show that an alternative version of the one-sector calibration with a more persistent and less volatile z process can create a much larger cyclical amplification of the unemployment rate and a stronger Beveridge curve, but at the cost of missing many of the unemployment duration targets, including the survival functions and the unemployment duration distributions for all workers and by age groups, as well as the cyclicality of the unemployment duration distribution. Thus, the one-sector version of our model appears unable to reconcile aggregate fluctuations with individual level patterns in unemployment outcomes. This trade-off disappears once unemployed workers are allowed to search across occupations, as the z^r cutoffs create narrower and more cyclically sensitive areas of inactions for each x_h.

5.2 Occupation Heterogeneity and Cyclical Unemployment

We now show that the same mechanisms described above hold within each task-based category but their strength varies across these occupational groups. Consequently, unemployed workers face different unemployment outcomes that depend also on the identity of the occupation. Both the long-run and cyclical dimensions of occupation-wide productivity differences are relevant. To understand the former, column 5 in Table 7 shows the contribution of workers who switched occupations through spells of unemployment in changing the observed sizes of the task-based categories in our calibration. This is compared to the contribution of the model's exogenous entry and exit process as captured by d and ψ_o (column 4 "Entrants"), such that for each task-based category the two values add up to the change in the employment stock. The calibration shows that NRM occupations increased in size due to more workers switching to these occupations through unemployment than away from them. In contrast, RM and RC decrease in size as more workers switching occupations through unemployment

Table 7: Role Unemployment in the Changing Size of Occupations

Task-Based	Distributions			Model Decomposition of Distribution Change			
	Initial	End Distribution		Entrants	Occ. Mob through Unemployment		
Occupational Categories	Distribution	Data	Model	All Qtrs	All Qtrs	Qtrs $u < u^{median}$	Qtrs $u \geq u^{median}$
Non-routine Cognitive	0.224	0.329	0.337	0.133	-0.020	-0.011	- 0.009
Routine Cognitive	0.292	0.258	0.246	-0.019	-0.027	-0.009	- 0.018
Non-routine Manual	0.226	0.260	0.260	-0.036	0.070	0.025	0.045
Routine Manual	0.258	0.154	0.157	-0.067	-0.034	-0.008	- 0.026

move away from these occupations than to them.

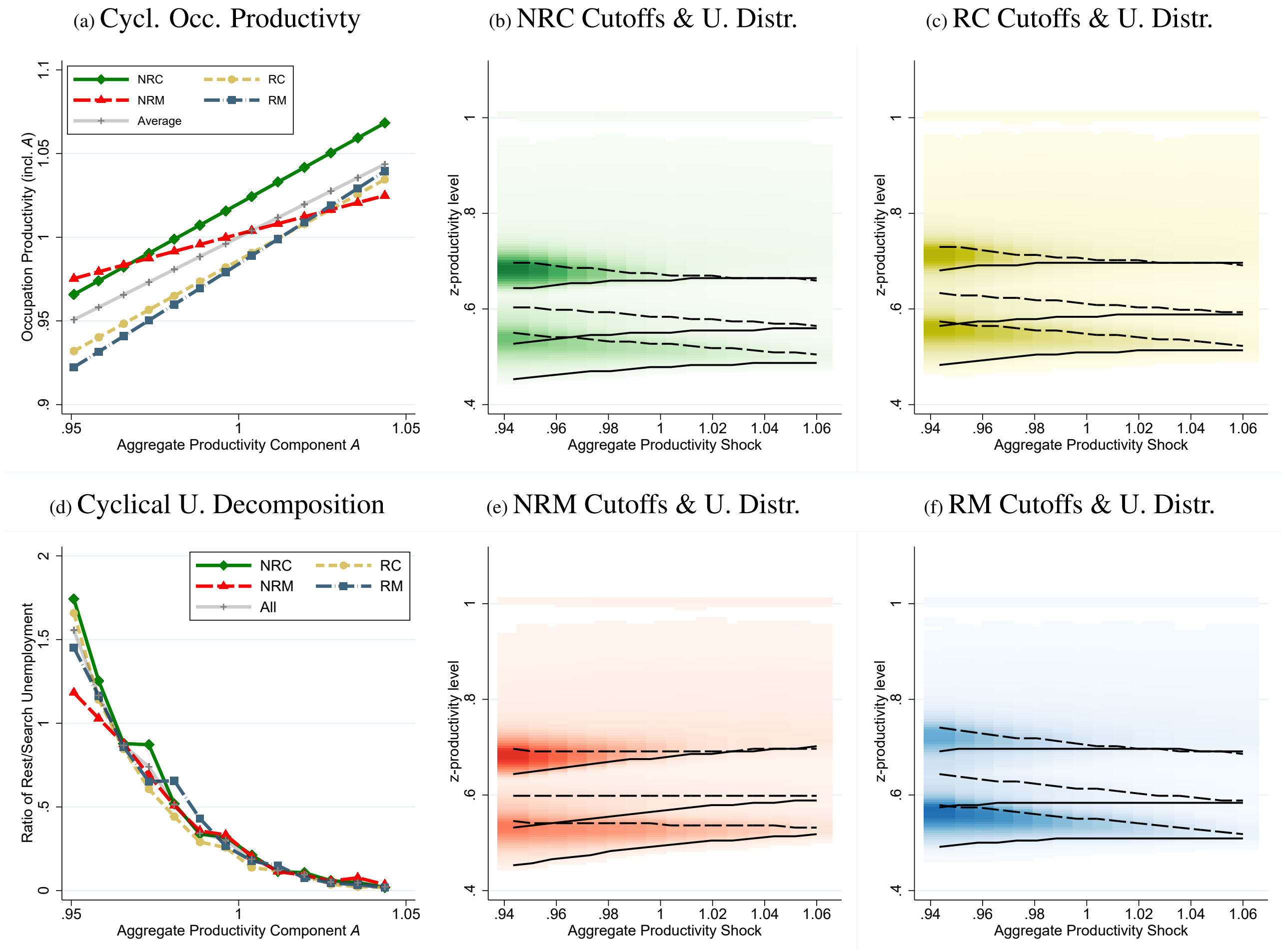
The last two columns of Table 7 show the contribution of mobility through unemployment separately by periods of high and low unemployment, where we categorise these periods by comparing the HP-filtered unemployment rate to its median. We observe that it is during recessions that mobility through unemployment particularly accelerates the changing size of NRM and RM occupations, representing about two-thirds and three-quarters of the total contribution of this channel, respectively. Jaimovich and Siu (2020) already documented the importance of recessions in changing the size of routine occupations. Here we show that the net mobility patterns described in Section 2 together with the endogenous response in unemployment yield precisely such a pattern within our model. Figure 11 illustrates the mechanism behind this. Figure 11a shows the levels and cyclicalities of the estimated occupation-wide productivities for the range of A. Reflecting the estimated values of ϵ_o, it shows that RM and RC occupations are very adversely affected in recessions, but catch up with the average in expansions. In contrast, NRM occupations are the least attractive in expansions for the unemployed but become more attractive in recessions. NRC occupations are consistently above average over the cycle (more so in expansions).

Figures 11b, 11c, 11e, 11f show that these different cyclical productivities result in different separation and reallocation cutoffs. Although their levels are not that different across task-based categories, in RM occupations the separation cutoffs decreases more steeply, while the reallocation cutoffs are nearly horizontal. In NRM occupations the separation cutoffs are nearly horizontal and the reallocation cutoffs are strongly upward-sloping. This implies that in recessions job separations are more prominent in RM than in NRM occupations.

Despite the differences in slopes, all task-based categories exhibit cutoffs with the $z^s > z^r$ property. Further, the distance between these cutoffs creates areas of inaction that increase in recessions, "trapping" a larger mass of unemployed workers in rest unemployment episodes, and narrow in expansions as workers increase their occupational mobility or re-gain employment in their pre-separation occupations. Figure 11d shows that as a result rest unemployment episodes become more common than search unemployment episodes in recessions within each task-based category, but as the economy recovers search unemployment episodes become the most common ones.

The observed countercyclical net mobility patterns then occur for mainly two reasons: (i) a differential cyclical response in the outflows across task-based categories, such that some task-based categories shed more workers during recessions relative to the average; and (ii) a differential cyclical

Figure 11: Heterogeneity across Occupation across the Cycle



response in the inflows, such that those workers who have decided to change occupations choose their destination task-based category differently in recessions than in expansions. The widening of the area of inactions as A decreases implies that overall occupational mobility falls during recessions and rises in expansions in all task-based categories. However, the differential responses in occupation-wide productivities across the business cycle depicted in Figure 11a imply that the decrease in outflows is stronger in NRM occupations and weaker in RM occupations relative to the average during recessions, as observed in the data. At the same time, Table 3 shows that the model is also able to reproduce the shift in the inflow distribution towards RM and away from NRM occupations that occurs in recessions.

6 Conclusions

In this paper we show that there is no tension between the cyclical behavior of individual unemployment outcomes, procyclical gross occupational mobility and countercyclical net mobility through unemployment. While individual outcomes are to a large extent driven by the interaction between worker-occupation idiosyncratic productivities and aggregate productivity, net mobility is affected by occupation-wide productivity differences (with different cyclical sensitivities) and unemployed workers' differential responses to these. Further, given that both the number of unemployment spells and the net mobility increase in recessions, transitions through unemployment play a meaningful role in shaping the changing size of RM, RC and NRM occupations.

The cyclically varying occupational mobility decisions imply that the nature of unemployment changes over the business cycle. During expansions (and mild recessions) the typical worker is not able to find jobs that are currently available to him due to standard search frictions and search unemployment becomes the main source of aggregate unemployment. In contrast, as recessions get stronger the typical worker is not able to find jobs because there are no jobs posted for him. In this case, rest or wait unemployment becomes the main source of aggregate unemployment. We show that for these dynamics to translate into large cyclical changes in aggregate unemployment and its duration distribution it is crucial to allow for endogenous occupational mobility in a way that is consistent with the observed mobility patterns of the unemployed.

The concept of rest unemployment is closely related to that of mismatch, stock-flow and rationing unemployment. Shimer (2007), for example, defines mismatch unemployment as those workers who remain attached to a local labor market even though there are currently no jobs for them. In stock-flow matching models, as in Coles and Smith (1998), unemployed workers in the stock wait for new jobs to arrive, as existing vacancies do not offer suitable employment opportunities. In Michaillat (2012) rationing unemployment occurs because workers are currently unproductive and no jobs are posted for them. As conditions improve, they become productive and employable once again. A key difference with all these models, is that here workers in rest unemployment episodes always have the option of looking for jobs in alternative occupations. Crucially, the occupational mobility decision changes over the cycle, with a larger proportion of workers deciding not to use this option in recessions.

Throughout, our analysis we have considered workers who are currently in a rest unemployment episodes as part of the labor force, still searching and expecting a positive job finding probability in the near future. Episodes of rest unemployment, however, could also conceptually be extended to incorporate marginally attached workers. In terms of the occupational mobility patterns, Section 2 shows that our analysis is robust to introducing periods of non-participation within workers' jobless spells. Appendix C.2.3 shows that considering non-participation periods in our targeted statistics does not alter the quantitative performance of our model. These exercises suggest that our results are then robust to inclusion of the marginally attached.

Although other models have been successful in replicating some of the cyclical unemployment patterns described here, Bils et al. (2011) argue that these models would typically have difficulty in jointly explaining the observed cyclicality of the aggregate unemployment rate and generating realistic dispersion in wage growth. This should not be an issue in our framework. As shown in Sections 4 and 5 our calibration generates the observed cyclicality of unemployment together with a realistic amount of wage dispersion as measured by Hornstein et al. (2011) Mm ratio. In this paper we have emphasised labor market flows pertaining to the unemployed, but extending the analysis to include heterogeneity in firm-worker matches and on-the-job search would allows us to study the cyclical relationship between wages, occupational mobility, unemployment fluctuations. We leave this topic for future research.

References

[1] Abraham, K. and L. Katz. 1986. "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances?". *Journal of Political Economy*, 94(3): 507-522.

[2] Abowd, J. and A. Zellner. 1985. "Estimating Gross Labor-Force Flows". *Journal of Business & Economic Statistics*, 3(3): 254-283.

[3] Ahn, H. J. and J. D. Hamilton. 2018. "Heterogeneity and Unemployment Dynamics". *Journal of Business & Economic Statistics*, forthcoming.

[4] Alvarez, F. and R. Shimer. 2011. "Search and Rest Unemployment". *Econometrica*, 79(1): 75-122.

[5] Alvarez, F. and R. Shimer. 2012. "Unemployment and Human Capital". Mimeo, University of Chicago, USA.

[6] Autor, D. H., F. Levy, R. J. Murnane. 2003. "The skill Content of Recent Technological Change: An Empirical Exploration". *Quarterly Journal of Economics*, Vol. 116(4): 1279-1333.

[7] Belot, M., P. Kircher and P. Muller. 2018. "Providing Advice to Job Seekers at Low Cost". *Review of Economic Studies*, forthcoming.

[8] Bils, M., Y. Chang and S. Kim. 2012. "Comparative Advantage and Unemployment Worker". *Journal of Monetary Economics*, 59: 150-165.

[9] Bils, M., Y. Chang and S. Kim. 2011. "Worker Heterogeneity and Endogenous Separations in a Matching Model of Unemployment Fluctuations". *American Economic Journal: Macroeconomics*, 3(1): 128-154.

[10] Carrillo-Tudela, C., B. Hobijn, P. She and L. Visschers. 2016. "The Extent and Cyclicality of Career Changes: Evidence for the U.K." *European Economic Review*, Vol. 84: 18-41.

[11] Carrillo-Tudela, C., B. Hobijn, and L. Visschers. 2014. "Career Changes Decline During Recessions". *FRBSF Economic Letters*, No. 2014-09.

[12] Chassamboulli, A. 2013. "Labor-Market Volatility in a Matching Model with Worker Heterogeneity and Endogenous Separations". *Labour Economics*, 24: 217-229.

[13] Coles, M. G. and E. Smith. 1998. "Marketplaces and Matching". *International Economic Review*, 39(1): 239-254.

[14] Cortes, M., N. Jaimovich, C. J. Nekarda and H. E. Siu. 2020. "The Micro and Macro Disappearing Routine Jobs: A Flow Approach". *Labour Economics*, forthcoming.

[15] Dvorkin, M. 2014. "Sectoral Shocks, Reallocation and Unemployment in Competitive Labor Markets ". Mimeo. Yale University, USA.

[16] Faberman, J. and M. Kudlyak. 2019. "The Intensity of Job Search and Search Duration". *American Economic Journal: Macroeconomics*, Vol. 11(3): 327-357.

[17] Fujita, S. and G. Moscarini. 2017. "Recall and Unemployment". *American Economic Review*, Vol. 102(7): 3875-3916.

[18] Gouge, R., and I. King. 1997. “A Competitive Theory of Employment Dynamics”. *Review of Economic Studies*, 64(1): 1-22.

[19] Hagedorn, M. and I. Manovskii. 2008.“The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited”. *American Economic Review*, 98(4): 1692-1706.

[20] Hamilton, J. 1988. “A Neoclassical Model of Unemployment and the Business Cycle”. *Journal of Political Economy*, 96(3): 593-617.

[21] Hall, R. and P. Milgrom. 2008. “The Limited Influence of Unemployment on the Wage Bargain”. *American Economic Review*, 98(4): 1653-1674.

[22] Hausman, J., J. Abrevaya, and F.M. Scott-Morton. 1998. “Misclassification of the Dependent Variable in a Discrete-Response Setting”, *Journal of Econometrics*, 87: 239-269.

[23] Hornstein, A. 2013. “Accounting for Unemployment: The Short and the Long of It”. Working Paper Series, Federal Reserve Bank of Richmond, WP 12-07.

[24] Hornstein, A., P. Krusell and G. L. Violante. 2011. “Frictional Wage Dispersion in Search Models: A Quantitative Assessment”, *American Economic Review*, 101(7): 2873-2898.

[25] Huckfeldt, C. 2016. “Understanding the Scarring Effects of Recessions”. Mimeo, Department of Economics, Cornell University, US.

[26] Jaimovich, N. and H. Siu. 2020. “The Trend is the Cycle: Job Polarization and Jobless Recoveries”. *Review of Economic and Statistics*, forthcoming.

[27] Jovanovic, B. and R. Moffitt. 1990. “An Estimate of a Sectoral Model of Labor Mobility”. *Journal of Political Economy*, 98(4): 827-852.

[28] Jovanovic, B. 1987. “Work, Rest and Search: Unemployment, Turnover, and the Cycle”. *Journal of Labor Economics*, 5(2): 131-148.

[29] Kambourov, G. and I. Manovskii. 2013. “A Cautionary Note on Using (March) Current Population Survey and Panel Study of Income Dynamics Data to Study Worker Mobility”, *Macroeconomic Dynamics*, 17(1): 172-194.

[30] Kambourov, G. and I. Manovskii. 2009a. “Occupational Mobility and Wage Inequality”. *Review of Economic Studies*, 76(2): 731-759.

[31] Kambourov, G. and I. Manovskii. 2009b. “Occupational Specificity of Human Capital”. *International Economic Review*, 50(1): 63-115.

[32] Kambourov, G. and I. Manovskii. 2008. “Rising Occupational and Industry Mobility in the United States: 1968-97”. *International Economic Review*, 49(1): 41-79.

[33] Keane, M. and K. Wolpin, 2011. “The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment”. *International Economic Review*, 42(4): 1051-1103.

[34] Kennan, J. and J. Walker. 2011. “The Expected Lifetime Income of Individual Migration Decisions”. *Econometrica*, 79(1): 211-251.

[35] Lilien, D. 1982. “Sectoral Shifts and Cyclical Unemployment”. *Journal of Political Economy*, 90(4): 777-793.

[36] Lucas, R. and E. Prescott. 1974. “Equilibrium Search and Unemployment”. *Journal of Economic Theory*, 7: 188-209.

[37] Mathiowetz, N. 1992. “Errors in Reports of Occupations”. *Public Opinion Quarterly*, 56: 332-135.

[38] Menzio, G. and S. Shi. 2010. “Block recursive equilibria for stochastic models of search on the job.” *Journal of Economic Theory*, 145(4): 1453-1494.

[39] Menzio, G. and S. Shi. 2011. “Efficient Search on the Job and the Business Cycle”. *Journal of Political Economy*, 119(3): 468-510.

[40] Michaillat, P. 2012. “Do Matching Frictions Explain Unemployment? Not in Bad Times”. *American Economic Review*, 102(4): 1721-1750.

[41] Mortensen, D. and C. Pissarides. 1994. “Job Creation and Job Destruction in the Theory of Unemployment”. *Review of Economic Studies*, 61(3): 397-415.

[42] Moscarini, G. and K. Thomsson. 2007. “Occupational and Job Mobility in the US”. *Scandinavian Journal of Economics*, 109(4): 807-836.

[43] Murtin, F. and JM. Robin. 2018. “Labor Market Reforms and Unemployment Dynamics”. *Labour Economics*, 50: 3-19.

[44] Mueller, A. I. 2017. “Separations, Sorting, and Cyclical Unemployment”. *American Economic Review*, 107(7): 2081-2107.

[45] Neal, D. 1999. “The Complexity of Job Mobility Among Young Men,” *Journal of Labor Economics*, Vol. 17(2): 237-524.

[46] Papageorgiou, T. 2018. “Worker Sorting and Agglomeration Economies”. Mimeo, Department of Economics, Boston College, US.

[47] Petrongolo, B. and C. Pissarides. 2001. “Looking into the Black Box: A Survey of the Matching Function”. *Journal of Economic Literature*, 39: 390-43.

[48] Pilossoph, Laura. 2013. “A Multisector Equilibrium Search Model of Labor Reallocation”. Mimeo, University of Chicago, USA.

[49] Poterba, J. and L. Summers. 1986. “Reporting Errors and Labor Market Dynamics”. *Econometrica*, 54(6): 1319-1338.

[50] Rogerson, R. 2005. “Sectoral Shocks, Specific Human Capital and Displaced Workers,” *Review of Economic Dynamics*, 8(1): 89-105.

[51] Rogerson, R. 1987. “An Equilibrium Model of Sectoral Reallocation,” *Journal of Political Economy*, 95(4): 824-834.

[52] Şahin, A., J. Song, G. Topa and G. L. Violante. 2014. “Mismatch Unemployment”. *American Economic Review*, 104(11): 3529-3564.

[53] Shimer, Robert. 2007. “Mismatch”. *American Economic Review*, 97(4): 1074-1101.

[54] Shimer, Robert. 2005. “The Cyclical Behavior of Equilibrium Unemployment and Vacancies”. *American Economic Review*, 95(1): 25-49.

[55] Veracierto, Marcelo. 2008. “On the Cyclical Behavior of Employment, Unemployment and Labor Force Participation”. *Journal of Monetary Economics*, 55: 1143-1157.

[56] Wiczer, David. 2015. “Long-term Unemployment: Attached and Mismatched?” Research Division, Working Paper Series, Federal Reserve Bank of St. Louis, WP 2015-042A.

APPENDIX

A The importance of coding error in occupational mobility

This appendix complements Section 2 of the paper. Here we quantify the extent to which coding errors affect the probability of an occupational change after a non-employment spell. This is done in two ways. (1) We exploit the change from independent to dependent interviewing that occurred across the 1985 and 1986 SIPP panels. This change of interviewing technique allows us to identify and estimate the probabilities that a worker's true occupation i gets miscoded as another occupation j. These probabilities are then used to derive a correction matrix that eliminates the impact of coding errors on the occupational mobility statistics obtained from the SIPP. (2) We then take advantage of the retrospective coding exercise done to occupational codes in the PSID. Retrospective coding improved the assignment of the occupational codes obtained during the 1970s, but did not affect the codes obtained during later years (see also Kambourov and Manovskii, 2008). We use probabilistic models to estimate the average effect retrospective coding had on the probability of an occupational change.

Across these data sets we obtain a very similar conclusion. Our correction method implies that on average 82% of the observed wave-to-wave occupational transitions after re-employment in the SIPP are genuine. In stark contrast, when pooling together workers who changed and did not change employers, we find that 40% of the observed changes among all workers are genuine. Similarly, retrospective coding in the PSID implies that 84% of the observed year-to-year occupational transitions after re-employment are genuine. When pooling together workers who changed and did not change employers, retrospective coding implies that only 44% are genuine. A key insight is that the same propensities to miscode occupations affect very differently the measured occupational change of employer movers and employer stayers.

A.1 A classification error model

In Section 2.1 of the main text we defined the garbling matrix $\mathbf{\Gamma}$, where its elements are the probabilities that an occupation i is miscoded as an occupation j, for all $i, j = 1, 2, ..., O$. We make three assumptions that will allows us to identify and estimate $\mathbf{\Gamma}$. (*A1*) *Independent classification errors*: conditional on the true occupation, the realization of the occupational code is independent of worker history, worker characteristics or time. (*A2*) *"Detailed balance" in miscoding*: $\mathbf{diag}(\mathbf{c})\mathbf{\Gamma}$ is symmetric, where $\mathbf{c}$ is a Ox1 vector that describes the distribution of workers across occupations and $\mathbf{diag}(\mathbf{c})$ is the diagonal matrix of $\mathbf{c}$. (*A3*) *Strict diagonal dominance*: $\mathbf{\Gamma}$ is strictly diagonally dominant in that $\gamma_{ii} > 0.5$ for all $i = 1, 2, ..., O$.

The first assumption is consistent with coder error. In the standard practise of independent interviewing, professional coders base their coding on the verbatim description of the reported work activities without taking into account the respondents' demographic characteristics or earlier work history.[27] Errors introduced by the respondents, however, could be correlated with their characteristics. Assumption *A1* implies that errors in the individuals' verbatim responses are fully captured by the nature of the job these individuals are performing and hence only depend on their *true* occupation.[28] The second assumption implies that coding error does not artificially change the size of occupations. This assumption is known as "detailed balance" and implies that the number of workers whose true occupation i gets mistakenly coded as j is the same as the number of workers whose true occupation j gets mistakenly coded as i. The third assumption implies that it is more likely to correctly code a given occupation i than to miscoded it. The converse would imply occupational mobility rates that are of a magnitude inconsistent with our data. Indeed, we derive an upper bound on code error directly from the data and find that *A3* is satisfied.

To estimate $\mathbf{\Gamma}$ we exploit the change of survey design between the 1985 and 1986 SIPP panels. Until the 1985 panel the SIPP used independent interviewing for all workers: in each wave all workers were asked to describe their job anew, without reference to answers given at an earlier date. Subsequently, a coder would consider that wave's verbatim descriptions and allocate occupational codes. This practise is known to generate occupational coding errors. In the 1986 panel, instead, the practise changed to one of dependent interviewing (see Lynn and Sala, 2007, Jäckle, 2009, and Jäckle and Eckman, 2019). Respondents were only asked "independently" to describe their occupation if they reported a change in employer or if they reported a change in their main activities without an employer change. If respondents declared no change in employer *and* in their main activities, the occupational code assigned to the respondent in the previous wave is carried forward.

There is an important literature that investigates classification error models. Magnac and Visser (1999) identify two main branches. The first one uses assumptions on the measurement error process and auxiliary data on the error rates, where it is assumed that the error rates can be directly observe from the auxiliary data (see Abowd and Zellner, 1985, Poterba and Summers, 1986, and Magnac and Visser, 1999, among others). The second one does not rely on auxiliary data on error rates, but estimates parametric models in the presence of misclassified data, where these models can be reduced form statistical models (see Hausman et al. 1998, among others) or structural economic models (see Kean and Wolpin, 2001, Sullivan, 2009, Roys and Taber, 2017, among others).[29]

Our classification error model builds on Abowd and Zellner (1985) and Poterba and Summer (1986) in that it uses a garbling matrix that captures the errors made in classifying workers. These authors investigate the misclassification of individuals' employment status in the CPS and use reported employment status at the original interview date, at the re-interviewing date (which occurred one week after the original interview) and the reconciliation information provided by the CPS to directly

[27] For example, during the 1980s and 1990s independent occupational coding in the PSID was done without reference to respondents' characteristics or their work history. However, this information was used in the retrospective coding exercise done to the 1970s occupational codes.

[28] See Mathiowetz (1992) for evidence of the importance of respondent error.

[29] Our estimation of coding errors differs from Keane and Wolpin (2001), Sullivan (2009) and Roys and Taber (2017) in that it does not rely on the observed optimal choices of economic agents for identification. Rather, miscoding can be estimated using the aforementioned occupational transitions alone (without requiring further information e.g. on wages). Therefore, it is simple to implement without the need for simulation methods and thereby reduces the overall computational burden when estimating our model.

observe the garbling matrix. Their key assumption is that the reconcile information provides the true individuals' labor market status. In contrast, we do not have auxiliary information that allows us to directly recover the garbling matrix. Our challenge is to estimate the garbling matrix. In what follows we show that given assumptions A1-A3, $\mathbf{\Gamma}$ is identified and can be estimated from the 1985 and 1986 SIPP panels.

To identify $\mathbf{\Gamma}$ it is important to note that during February 1986 to April 1987, the 1985 and 1986 panels overlap, representing the *same* population under different survey designs. The identification theory we develop in the next section refers to this population. We then show how to consistently estimate $\mathbf{\Gamma}$ using the population samples.

A.1.1 Identification of $\mathbf{\Gamma}$

Consider the population represented by 1985/86 panels during the overlapping period and divide it into two groups of individuals across consecutive interviews by whether or not they changed employer or activity. Label those workers who stayed with their employers in both interviews and did not change activity as "employer/activity stayers". By design this group *only* contains true occupational stayers. Similarly, label those workers who changed employers or changed activity within their employers as "employer/activity changers". By design this group contains all true occupational movers and the set of true occupational stayers who changed employers.

Suppose that we were to subject the employer/activity stayers in this population to dependent interviewing as applied in the 1986 panel. Let $\mathbf{c_s}$ denote the Ox1 vector that describes their *true* distribution across occupations and let $\mathbf{M_s} = \mathbf{diag}(\mathbf{c_s})$. Let $\mathbf{c_s^D}$ denote the Ox1 vector that describes their *observed* distribution across occupations under dependent interviewing and let $\mathbf{M_s^D} = \mathbf{diag}(\mathbf{c_s^D})$. Note that $\mathbf{c_s^D} = \mathbf{\Gamma}'\mathbf{M_s}\overrightarrow{\mathbf{1}}$, where $\overrightarrow{\mathbf{1}}$ describes a vector of ones. $\mathbf{M_s}$ is pre-multiplied by $\mathbf{\Gamma}'$ as true occupations would have been miscoded in the first of the two consecutive interviews. Assumption *A2* implies that $\mathbf{c_s^D} = \mathbf{diag}(\mathbf{c_s})\mathbf{\Gamma}\overrightarrow{\mathbf{1}} = \mathbf{c_s}$ and hence $\mathbf{M_s^D} = \mathbf{M_s}$.

Next suppose that instead we were to subject the employer/activity stayers in this population to independent interviewing as applied in the 1985 panel. Let $\mathbf{M_s^I}$ denote the matrix that contains these workers' *observed* occupational transition *flows* under independent interviewing. In this case $\mathbf{M_s^I} = \mathbf{\Gamma}'\mathbf{M_s}\mathbf{\Gamma}$. Here $\mathbf{M_s}$ is pre-multiplied by $\mathbf{\Gamma}'$ and post-multiplied by $\mathbf{\Gamma}$ to take into account that the observed occupations of origin and destination would be subject to coding error.

Let $\mathbf{M_m}$ denote the matrix that contains the *true* occupational transition *flows* of employer/activity changers in this population. The diagonal of $\mathbf{M_m}$ describes the distribution of true occupational stayers across occupations among employer/activity changers. The off-diagonal elements contain the flows of all true occupational movers. Under independent interviewing $\mathbf{M_m^I} = \mathbf{\Gamma}'\mathbf{M_m}\mathbf{\Gamma}$. Once again $\mathbf{M_m}$ is pre-multiplied by $\mathbf{\Gamma}'$ and post-multiplied by $\mathbf{\Gamma}$ as the observed occupations of origin and destination would be subject to coding error.

Letting $\mathbf{M^I} = \mathbf{M_m^I} + \mathbf{M_s^I}$ denote the matrix that contains the aggregate occupational transition flows across two interview dates under independent interviewing, it follows that $\mathbf{M_s^I} = \mathbf{M^I} - \mathbf{M_m^I} = \mathbf{\Gamma}'\mathbf{M_s}\mathbf{\Gamma}$. By virtue of the symmetry of $\mathbf{M_s}$ and assumption *A2*, $\mathbf{M_s}\mathbf{\Gamma} = \mathbf{\Gamma}'\mathbf{M_s}' = \mathbf{\Gamma}'\mathbf{M_s}$. Substituting back yields $\mathbf{M}_\mathrm{s}^\mathbf{I} = \mathbf{M}_\mathrm{s}\mathbf{\Gamma}\mathbf{\Gamma}$. Next note that $\mathbf{M}_\mathrm{s}^\mathbf{I} = \mathbf{M}_\mathrm{s}\mathbf{T}_\mathrm{s}^\mathbf{I}$, where $\mathbf{T}_\mathrm{s}^\mathbf{I}$ is the occupational transition probability matrix of the employer/activity stayers in this population *observed* under independent interviewing. Substitution yields $\mathbf{M}_\mathrm{s}\mathbf{T}_\mathrm{s}^\mathbf{I} = \mathbf{M}_\mathrm{s}\mathbf{\Gamma}\mathbf{\Gamma}$. Multiply both sides by $\mathbf{M}_\mathrm{s}^{-1}$, which exists as long as all the diagonal elements of $\mathbf{M}_\mathrm{s}$ are non-zero, yields the key relationship we exploit to estimate $\mathbf{\Gamma}$,

$$\mathbf{T}_\mathrm{s}^\mathbf{I} = \mathbf{\Gamma}\mathbf{\Gamma}. \tag{7}$$

To use this equation we first need to show that it implies a unique solution for $\mathbf{\Gamma}$. Towards this result, we now establish that $\mathbf{\Gamma}$ and $\mathbf{T}_\mathrm{s}^\mathbf{I}$ are diagonalizable. For the latter it is useful to interpret the coding error process described above as a Markov chain such that $\mathbf{\Gamma}$ is the one-step probability matrix associated with this process.

Lemma A.1: *Assumptions A2 and A3 imply that* $\mathbf{\Gamma}$ *and* $\mathbf{T}_\mathrm{s}^\mathbf{I}$ *are diagonalizable.*

Proof. First note that without loss of generality we can consider the one-step probability matrix $\mathbf{\Gamma}$ to be irreducible. To show this suppose that $\mathbf{\Gamma}$ was not irreducible, we can (without loss of generality) apply a permutation matrix to re-order occupations in $\mathbf{\Gamma}$ and create a block-diagonal $\mathbf{\Gamma}'$, where each block is irreducible and can be considered in isolation. Given *A3*, it follows directly that $\mathbf{\Gamma}$ is aperiodic. Further, assumption *A2* implies that $\mathbf{c}_\mathrm{s}$ is a stationary distribution of $\mathbf{\Gamma}$. The fundamental theorem of Markov chains then implies that $\mathbf{c}_\mathrm{s}$ is the *unique* stationary distribution of $\mathbf{\Gamma}$. Assumption *A2* also implies that the Markov chain characterised by $\mathbf{\Gamma}$ is reversible with respect to $\mathbf{c}_\mathrm{s}$. This means that $\mathbf{\Gamma}$ is similar (in matrix sense) to a symmetric matrix $\mathbf{G}$ such that $\mathbf{G} = \mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}})\ \mathbf{\Gamma}\ \mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}})^{-1}$. By symmetry, $\mathbf{G}$ is orthogonally diagonalizable by $\mathbf{Q}\mathbf{\Delta}\mathbf{Q}^{-1}$, where the diagonal matrix $\mathbf{\Delta}$ contains the associated (real) eigenvalues and $\mathbf{Q}$ is the orthogonal matrix of associated (normalized) eigenvectors. It then follows that $\mathbf{\Gamma}$ is diagonalizable as well. Further, $\mathbf{T}_\mathrm{s}^\mathbf{I} = \mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}})^{-1}\ \mathbf{G}\ \mathbf{G}\ \mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}}) = \mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}})^{-1}\ \mathbf{Q}\ \mathbf{\Delta}^2\ \mathbf{Q}^{-1}\ \mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}})$, and hence $\mathbf{T}_\mathrm{s}^\mathbf{I}$ is also orthogonally diagonalizable, with a root of $\mathbf{P}\mathbf{\Lambda}^{\mathbf{0.5}}\mathbf{P}^{-1}$, where $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues of $\mathbf{T}_\mathrm{s}^\mathbf{I}$, and $\mathbf{P}$ the associated orthogonal matrix with eigenvectors of $\mathbf{T}_\mathrm{s}^\mathbf{I}$. □

In general one cannot guarantee the uniqueness, or even existence, of a transition matrix that is the (nth) root of another transition matrix (see Higham and Lin, 2011). Here, however, existence is obtained by construction: $\mathbf{T}_\mathrm{s}$ is constructed from $\mathbf{\Gamma}$, and in reverse, we can find its roots. The next result shows that $\mathbf{T}_\mathrm{s}$ has a unique root satisfying assumptions *A2* and *A3*.

Proposition A.1: $\mathbf{\Gamma}$ *is the unique solution to* $\mathbf{T}_\mathrm{s}^\mathbf{I} = \mathbf{\Gamma}\,\mathbf{\Gamma}$ *that satisfies assumptions A2 and A3. It is given by* $\mathbf{P}\mathbf{\Lambda}^{0.5}\mathbf{P}^{-1}$*, where* $\mathbf{\Lambda}$ *is the diagonal matrix with eigenvalues of* $\mathbf{T}_\mathrm{s}^\mathbf{I}$, $0 < \lambda_i \leq 1$*, and* $\mathbf{P}$ *is the orthogonal matrix with the associated (normalized) eigenvectors.*

Proof. Following from the proof of Lemma A.1, a root of $\mathbf{T}_\mathrm{s}^\mathbf{I}$ is given by $\mathbf{P}\mathbf{\Lambda}^{0.5}\mathbf{P}^{-1}$, where $\mathbf{\Lambda}$ is the diagonal matrix with eigenvalues of $\mathbf{T}_\mathrm{s}^\mathbf{I}$ and $\mathbf{P}$ is the orthogonal matrix with the associated (normalized) eigenvectors. Since *A3* implies $\mathbf{\Gamma}$ is strictly diagonally dominant, it follows that the determinant of all its leading principal minors are positive. Moreover, under the similarity transform by pre-/post-multiplication with the diagonal matrices $\mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}}), \mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}})^{-1}$, the determinant of all principals minors of the symmetric matrix $\mathbf{G} = \mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}})\ \mathbf{\Gamma}\ \mathbf{diag}(\sqrt{\mathbf{c}_\mathrm{s}})^{-1}$ are positive as well. Hence $\mathbf{G}$ is a symmetric positive definite matrix (with all eigenvalues between 0 and 1, as has $\mathbf{\Gamma}$). It follows that $\mathbf{G}\,\mathbf{G} = \mathbf{S}$ is also positive definite, and $\mathbf{T_s^I} = \mathbf{diag}(\sqrt{\mathbf{c_s}})^{-1}\,\mathbf{S}\,\mathbf{diag}(\sqrt{\mathbf{c_s}})$ is positive definite in the sense that $\mathbf{v}'\,\mathbf{T_s^I}\,\mathbf{v} > 0$ for all $\mathbf{v} \neq \mathbf{0}$, while also all eigenvalues of $\mathbf{T_s^I}$ will be between 0 and 1.

To show the uniqueness of the root of $\mathbf{T_s^I}$ suppose (towards a contradiction) that there exists two different roots $\mathbf{\Gamma}$ and $\mathbf{\Upsilon}$ such that each are similar (in matrix sense), with the same transform involving $\mathbf{diag}(\sqrt{\mathbf{c_s}})$, to different symmetric positive definite matrices $\mathbf{G}$ and $\mathbf{Y}$, where $\mathbf{G}\,\mathbf{G} = \mathbf{S}$ and $\mathbf{Y}\,\mathbf{Y} = \mathbf{S}$. Both $\mathbf{G}$ and $\mathbf{Y}$ are diagonalizable, and have the square roots of the eigenvalues of $\mathbf{S}$ on the diagonal. Given that the squares of the eigenvalues need to coincide with the eigenvalues of $\mathbf{S}$ and assumptions *A2* and *A3* imply that all eigenvalues must be between 0 and 1, without loss of generality we can consider both diagonalizations to have the same diagonal matrix $\mathbf{\Delta}$, where $\mathbf{\Delta}$ is the diagonal matrix of eigenvalues of $\mathbf{T_s^I}$ and these eigenvalues are ordered using a permutation-similarity transform with the appropriate permutation matrices. Let $\mathbf{G} = \mathbf{H\Delta H^{-1}}$ and $\mathbf{Y} = \mathbf{K\Delta K^{-1}}$. Then, it follows that $\mathbf{K^{-1}H\Delta^2H^{-1}K} = \mathbf{\Delta^2}$ and since $\mathbf{K^{-1}H}$ and $\mathbf{\Delta^2}$ commute, implies that $\mathbf{K^{-1}H}$ is a block-diagonal matrix with the size of the blocks corresponding to the multiplicity of squared eigenvalues. Again, since all eigenvalues of $\mathbf{\Delta}$ are positive, this equals the multiplicity of the eigenvalues δ_i itself. But then it must be true that $\mathbf{K^{-1}H\Delta H^{-1}K} = \mathbf{\Delta}$. Then, $\mathbf{G} = \mathbf{H}\,\mathbf{\Delta}\,\mathbf{H^{-1}} = \mathbf{KK^{-1}H}\,\mathbf{\Delta}\,\mathbf{H^{-1}KK^{-1}} = \mathbf{K}\,\mathbf{\Delta}\,\mathbf{K^{-1}} = \mathbf{Y}$ which leads to a contradiction. □

The above results imply that under assumptions *A2* and *A3*, $\mathbf{\Gamma}$ is uniquely identified from the transition matrix of true occupational stayers under independent interviewing, $\mathbf{T_s^I}$.

A.1.2 Estimation of $\mathbf{\Gamma}$

The next lemma provides an intermediate step towards estimating $\mathbf{\Gamma}$. For this purpose let $PDT(.)$ denote the space of transition matrices that are similar, in the matrix sense, to positive definite matrices.

Lemma A.2: *The function $f : PDT(\mathbb{R}^{O\times O}) \rightarrow PDT(\mathbb{R}^{O\times O})$ given by $f(\mathbf{T}) = \mathbf{T^{0.5}}$ exists and is continuous with $f(\mathbf{T_s^I}) = \mathbf{\Gamma}$ in the spectral matrix norm.*

Proof. Existence follows from Lemma A.1 and Proposition A.1. To establish continuity of the mapping, we follow Horn and Johnson (1990). Let $\mathbf{T_1}$ and $\mathbf{T_2}$ be any two transition matrices in PDT and let $\mathbf{U_1}$ and $\mathbf{U_2}$ be two symmetric positive definite matrices constructed as $\mathbf{U_1} = \mathbf{diag}(\sqrt{\mathbf{c_1}})\,\mathbf{T_1}\,\mathbf{diag}(\sqrt{\mathbf{c_1}})^{-1}$ and $\mathbf{U_2} = \mathbf{diag}(\sqrt{\mathbf{c_2}})\,\mathbf{T_2}\,\mathbf{diag}(\sqrt{\mathbf{c_2}})^{-1}$, where $\mathbf{c_1}$ and $\mathbf{c_2}$ are the unique stationary distributions associated with $\mathbf{T_1}$ and $\mathbf{T_2}$, respectively. We want to show that if $\mathbf{U_1} \rightarrow \mathbf{U_2}$, then $\mathbf{U_1^{0.5}} \rightarrow \mathbf{U_2^{0.5}}$. First, note that $\|\mathbf{U_1} - \mathbf{U_2}\|_2 = \|\mathbf{U_1^{0.5}}(\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}}) + (\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}})\mathbf{U_2^{0.5}}\|_2 \geq |\mathbf{x}'\mathbf{U_1^{0.5}}(\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}})\mathbf{x} + \mathbf{x}'(\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}})\mathbf{U_2^{0.5}}\mathbf{x}|$, where $\mathbf{x}$ is any normalised vector. Assumptions *A2* and *A3* imply $\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}}$ exists and is a symmetric matrix. Let $|\lambda| = \rho(\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}})$ be the absolute value of the largest eigenvalue of $\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}}$ and let $\mathbf{z}$ be the normalized eigenvector associated with λ. Note that $\|\mathbf{U_1} - \mathbf{U_2}\|_2 = |\lambda|$ and $(\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}})\mathbf{z} = \lambda\mathbf{z}$. Then $|\mathbf{z}'\mathbf{U_1^{0.5}}(\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}})\mathbf{z} + \mathbf{z}'(\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}})\mathbf{U_2^{0.5}}\mathbf{z}| \geq |\lambda|\,|\lambda_{min}^{0.5}(\mathbf{U_1}) + \lambda_{min}^{0.5}(\mathbf{U_2})| = \|\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}}\|_2\,(\lambda_{min}^{0.5}(\mathbf{U_1}) + \lambda_{min}^{0.5}(\mathbf{U_2}))$, where $\lambda_{min}(\mathbf{U_1})$ denotes the smallest eigenvalue of $\mathbf{U_2}$, which is positive by virtue of assumptions *A2* and *A3*. Then choose a $\delta = \varepsilon\lambda_{min}^{0.5}(\mathbf{U_1})$. It follows that if $\|\mathbf{U_1} - \mathbf{U_2}\|_2 < \delta$, then $\|\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}}\|_2 \times$

$\frac{(\lambda^{0.5}_{min}(\mathbf{U_1})+\lambda^{0.5}_{min}(\mathbf{U_2}))}{\lambda^{0.5}_{min}(\mathbf{U_1})} < \varepsilon$, and therefore $\|\mathbf{U_1^{0.5}} - \mathbf{U_2^{0.5}}\|_2 < \varepsilon$, which establishes the desired continuity. From the fact that $\mathbf{U_1} \to \mathbf{U_2}$ implies $\mathbf{U_1^{0.5}} \to \mathbf{U_2^{0.5}}$, it then also follows that $f(\mathbf{T})$ is continuous. □

Let $\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}}$ denote the sample estimate of $\mathbf{T}^{\mathbf{I}}_{\mathbf{s}}$ and let $\hat{\mathbf{\Gamma}}$ be estimated by the root $(\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}})^{0.5} \in PDT(\mathbb{R}^{O\times O})$ such that $\hat{\mathbf{\Gamma}} = (\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}})^{\mathbf{0.5}} = \hat{\mathbf{P}}\hat{\mathbf{\Lambda}}^{\mathbf{0.5}}\hat{\mathbf{P}}^{\mathbf{-1}}$, where $\hat{\mathbf{\Lambda}}$ is the diagonal matrix with eigenvalues of $\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}}$, $0 < \hat{\lambda}_i^{0.5} \leq 1$ and $\hat{\mathbf{P}}$ the orthogonal matrix with the associated (normalized) eigenvectors. We then have the following result.

Proposition A.2: $\mathbf{\Gamma}$ *is consistently estimated from* $(\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}})^{0.5} \in PDT(\mathbb{R}^{O\times O})$ *such that* $\hat{\mathbf{\Gamma}} = (\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}})^{\mathbf{0.5}} = \hat{\mathbf{P}}\hat{\mathbf{\Lambda}}^{\mathbf{0.5}}\hat{\mathbf{P}}^{\mathbf{-1}}$. *That is,* $plim_{n\to\infty}\hat{\mathbf{\Gamma}} = \mathbf{\Gamma}$.

Proof. From Lemma A.1 and Proposition A.1 it follows that if we know $\mathbf{T}^{\mathbf{I}}_{\mathbf{s}}$, then we can find the unique $\mathbf{\Gamma}$ that underlies it, constructing it from the eigenvalues and eigenvectors of $\mathbf{T}^{\mathbf{I}}_{\mathbf{s}}$. To estimate $\mathbf{T}^{\mathbf{I}}_{\mathbf{s}}$ one can use the sample proportion $\hat{\mathrm{M}}_{\mathbf{ij}}/\sum_{k=1}^{O}\hat{\mathrm{M}}_{\mathbf{ik}}$ and note that this converges in probability to $(\mathbf{T}^{\mathbf{I}}_{\mathbf{s}})_{\mathbf{ij}}$ (see Anderson and Goodman, 1957; Billingsley 1961, thm 1.1-3.) for all occupations, given assumptions *A2* and *A3*. Hence, $plim_{n\to\infty}\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}} = \mathbf{T}^{\mathbf{I}}_{\mathbf{s}}$. Then, we derive $\hat{\mathbf{\Gamma}}$ from $\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}}$ according to $\hat{\mathbf{P}}\hat{\mathbf{\Lambda}}^{\mathbf{0.5}}\hat{\mathbf{P}}^{\mathbf{-1}}$, per Proposition A.1. By continuity of the mapping in Lemma A.2, it follows that $plim_{n\to\infty}\hat{\mathbf{\Gamma}} = \mathbf{\Gamma}$, and our estimator is consistent. □

Note that to identify and estimate $\mathbf{\Gamma}$ in the SIPP it is not sufficient to directly compare the aggregate occupational transition flows under independent interviewing with the aggregate occupational transition flows under dependent interviewing. To show this let $\mathrm{M}^{\mathbf{D}} = \mathrm{M}^{\mathbf{I}}_{\mathbf{m}} + \mathrm{M}^{\mathbf{D}}_{\mathbf{s}}$ denote the matrix that contains the aggregate occupational transition flows across two interview dates under dependent interviewing for employer/activity stayers and under independent interviewing for employer/activity movers. Subtracting $\mathrm{M}^{\mathbf{I}} = \mathrm{M}^{\mathbf{I}}_{\mathbf{m}} + \mathrm{M}^{\mathbf{I}}_{\mathbf{s}}$ from $\mathrm{M}^{\mathbf{D}}$ yields $\mathrm{M}^{\mathbf{D}}_{\mathbf{s}} - \mathrm{M}^{\mathbf{I}}_{\mathbf{s}} = \mathrm{M}_{\mathbf{s}} - \mathbf{\Gamma}'\mathrm{M}_{\mathbf{s}}\mathbf{\Gamma}$. Given the symmetry assumed in *A2*, the latter expression has $0.5n(n-1)$ exogenous variables on the LHS and $0.5n(n+1)$ unknowns (endogenous variables) on the RHS, leaving $\mathbf{\Gamma}$ (and $\mathrm{M}_{\mathbf{s}}$) unidentified.

In addition to $\mathrm{M}^{\mathbf{D}} - \mathrm{M}^{\mathbf{I}} = \mathrm{M}_{\mathbf{s}} - \mathbf{\Gamma}'\mathrm{M}_{\mathbf{s}}\mathbf{\Gamma}$ one can use $\mathrm{M}^{\mathbf{D}} = \mathbf{\Gamma}'\mathrm{M}_{\mathbf{m}}\mathbf{\Gamma} + \mathrm{M}_{\mathbf{s}}$, which contains the remainder information. When $\mathrm{M}_{\mathbf{m}}$ has mass on its diagonal, however, this additional system of equations has n^2 exogenous variables on the LHS and n^2 unknowns (arising from $\mathrm{M}_{\mathbf{m}}$) on the RHS. This implies that with the n unknowns remaining from $\mathrm{M}^{\mathbf{D}} - \mathrm{M}^{\mathbf{I}} = \mathrm{M}_{\mathbf{s}} - \mathbf{\Gamma}'\mathrm{M}_{\mathbf{s}}\mathbf{\Gamma}$, one is still unable to identify $\mathbf{\Gamma}$ and $\mathrm{M}_{\mathbf{s}}$.

Corollary A.1: *If* $\mathrm{M}_{\mathbf{m}}$ *has mass on its diagonal,* $\mathbf{\Gamma}$ *cannot be identified from* $\mathrm{M}^{\mathbf{I}}$ *and* $\mathrm{M}^{\mathbf{D}}$ *alone.*

The intuition behind this result is that by comparing aggregate occupational transition flows under dependent and independent interviewing, it is unclear how many workers are 'responsible' for the change in occupational mobility between $\mathrm{M}^{\mathbf{D}}$ and $\mathrm{M}^{\mathbf{I}}$. Only when the diagonal of $\mathrm{M}_{\mathbf{m}}$ contains exclusively zeros, identification could be resolved and one can recover $\mathrm{M}_{\mathbf{s}}$, $\mathbf{\Gamma}$ and $\mathrm{M}_{\mathbf{m}}$ as the number of equations equals the number of unknowns.[30] An implication of the above corollary is that interrupted time-series analysis that is based on the difference in occupational mobility at the time

[30] However, in the SIPP this case is empirically unreasonable as it requires that all employer/activity changers be true occupational movers.

of a switch from independent to dependent interviewing, does not identify the precise extent of the average coding error, but provides a downwards biased estimate.

To identify $\mathbf{\Gamma}$, however, Proposition A.2 implies that one can use the observed occupational transition flows of a sample of *true* occupational stayers that are subject to two rounds of independent interviewing. Some of these workers will appear as occupational stayers and some of them as occupational movers. Ideally, such a sample of workers should be isolated directly from the 1985 panel. Unfortunately, the questions on whether the individual changed activity or employer were only introduced in the 1986 panel, as a part of the switch to dependent interviewing. As a result, the 1985 panel by itself does not provide sufficient information to separate employer/activity stayers from employer/activity movers. Instead we use 1986 panel to estimate $\hat{\mathbf{M}}^{\mathbf{I}}_{\mathbf{m}}$. We can infer $\mathbf{M}^{\mathbf{I}}_{\mathbf{s}}$ indirectly by subtracting the observed occupational transition flow matrix $\hat{\mathbf{M}}^{\mathbf{I}}_{\mathbf{m}}$ in the 1986 panel from the observed occupational transition flow matrix $\hat{\mathbf{M}}^{\mathbf{I}}$ in the 1985 panel. This is possible as the 1986 panel refers to the same underlying population as the 1985 panel and separates the employer/activity changers, who are independently interviewed.

Corollary A.2: $\hat{\mathbf{\Gamma}}$ *is consistently estimated from* $\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}}$ *when the latter is estimated from* $\hat{\mathbf{M}}^{\mathbf{I}} - \hat{\mathbf{M}}^{\mathbf{I}}_{\mathbf{m}}$

This result is important to implement our approach. It follows as the population proportions underlying each cell of $\hat{\mathbf{M}}_{\mathbf{s}}$, the sample estimate of $\mathbf{M}_{\mathbf{s}}$, are consistently estimated. In turn, the latter follows from the standard central limit theorem for estimating proportions, which applies to $\hat{\mathbf{M}}^{\mathbf{I}}$, $\hat{\mathbf{M}}^{\mathbf{I}}_{\mathbf{m}}$ and its difference. Proposition A.2 then implies that $\hat{\mathbf{\Gamma}}$ is consistently estimated.

A.1.3 Implementation

To implement our correction method we take the overlapping period of the 1985/86 panels.[31] To increase the sample size we also use observations from the 1987 panel for the period between February 1987 and April 1987.[32] This panel has an identical setup to that of the 1986 panel (dependent interviewing and other relevant aspects) and is likewise representative of the population during the period of study.

Interviews throughout the SIPP are conducted every four months and collect information pertaining to the last four months, where these four months are considered to be a wave. We compare the reported occupational code of a worker in a given interview with the reported occupational code of that worker in the subsequent interview. An observation is therefore a pair of occupational codes, a reported 'source' and (potentially identical) 'destination' occupation. To keep comparability across interviews as clean as possible, and to focus on measuring occupations in the primary job, we only consider those workers who throughout the two waves stayed in full-time employment. We further restrict attention to those workers who reported having only one employer at any point in time, do not have imputed occupations, were not enrolled in school and were between 19-66 years old. These

[31] See Hill (1994) for a similar strategy and a further comparison of the 1985 and 1986 panel samples in terms of survey design, attrition rates and panel conditioning.

[32] To avoid seasonality effects we re-weight all observations such that each observation in a given month has the same weight as another observation in any other month.

restrictions yield 28,302 observations for the 1985 panel, 27,801 observations for the 1986 panel and 5,922 observations for the 1987 panel.

Table 8: Demographic characteristics - February 1986 to April 1987

	SIPP 1985	SIPP 1986	SIPP 1987	p-value (no difference)
Education				
less than high school	14.72	15.27	14.55	0.386
high school grad	38.10	37.36	36.80	0.361
some college	24.51	24.94	24.71	0.546
college degree	22.67	22.43	23.95	0.746
Age category				
19-24	12.29	12.62	12.90	0.458
25-29	16.72	16.15	16.36	0.357
30-34	15.84	15.40	16.00	0.512
35-39	15.02	15.30	13.99	0.806
40-44	11.65	11.50	11.80	0.804
45-49	9.10	8.87	9.22	0.667
50-54	7.72	7.90	8.25	0.600
55-59	7.07	7.31	7.01	0.600
60-64	4.20	4.64	4.05	0.221
Ethicity				
white	86.29	86.57	86.25	0.729
black	10.62	10.81	10.73	0.782
american indian, eskimo	0.49	0.62	0.42	0.401
asian or pacific islander	2.60	2.01	2.60	0.090
Other				
men	54.20	55.27	53.93	0.112
married	65.51	66.15	64.54	0.550
living in metro area	76.33	76.24	75.97	0.885

Workers aged 19-66, not enrolled in school, in two adjacent waves, measured in the first month of the current wave, with employment in one firm only in the previous wave, and employment in one (but possibly different) employer only in the wave that follows, without any self-employment, with un-imputed occupations reported in both waves. Person weights are used to scale observations per month within panel group (1985 versus 1986+1987).

Tables 8 and 9 show the demographic and occupational characteristics (based on the major occupations of the 1990 SOC), respectively, of the samples across the three panels. In the last column of each table we test characteristic-by-characteristic whether the proportion of workers with a given characteristic in the 1985 sample is statistically indistinguishable from the proportion of workers with the same characteristic in the pooled 1986/87 sample. Across all the characteristics analysed, we cannot reject at a 5% level that the proportions in the 1985 sample and the corresponding proportions in the 1986-87 sample are the same. With the exception of the proportion of Asian Americans and the proportion of workers whose source occupation is management, similarity cannot be rejected even at a 10% level. Although not shown here, we also cannot rejected at a 10% level that the proportion of workers across source and destination industries are the same when comparing the 1985 and 1986/87 samples. This analysis thus confirm that the observations used for our exercise are taken from the same underlying population.[33]

[33]To further rule out any meaningful impact from the observed differences in occupational distributions, we re-calculated all statistics after re-balancing the weights on the source/destination occupations to create identical occupa-

Table 9: Distribution of workers across occupations - February 1986 to April 1987

Distribution across source occupations (occupation code) (%)				
	SIPP 1985	SIPP 1986	SIPP 1987	p-value (no difference)
managing occupations	12.31	13.11	13.99	0.064
professional speciality	13.40	12.92	13.18	0.380
technicians and related support	3.85	3.89	4.08	0.829
sales occ.	9.74	9.98	9.86	0.599
admin support	18.67	18.17	18.26	0.366
services	10.96	11.66	11.19	0.166
farming/fish/logging	1.09	1.08	1.03	0.940
mechanics and repairers	4.87	4.47	4.69	0.231
construction and extractive	3.47	3.65	3.60	0.503
precision production	4.01	4.19	3.75	0.643
machine operators/assemblers	9.27	8.89	8.63	0.355
transportation and materials moving	4.63	4.34	4.31	0.340
laborers	3.73	3.66	3.42	0.714
Distribution across destination occupations (occupation code) (%)				
managing occupations	12.58	13.27	14.18	0.103
professional speciality	13.31	12.90	13.10	0.451
technicians and related support	3.82	3.92	4.01	0.693
sales occ.	9.76	9.89	9.77	0.797
admin support	18.53	18.20	18.20	0.546
services	10.96	11.57	11.19	0.229
farming/fish/logging	1.08	1.07	1.03	0.924
mechanics and repairers	4.87	4.46	4.73	0.222
construction and extractive	3.59	3.60	3.56	0.954
precision production	4.01	4.26	3.77	0.480
machine operators/assemblers	9.24	8.93	8.77	0.455
transportation and materials moving	4.62	4.31	4.33	0.310
laborers	3.62	3.62	3.36	0.906

Workers aged 19-66, not enrolled in school, in two adjacent waves, measured in the first month of the current wave, with employment in one firm only in the previous wave, and employment in one (but possibly different) firm only in the wave that follows, without any self-employment, with un-imputed occupations reported in both waves. Person weights are used to scale observations per month within panel group (1985 versus 1986+1987).

We estimate the occupational flows of employer/activity stayers by $\hat{\mathbf{M}}^{\mathbf{I},\mathbf{85}} - \hat{\mathbf{M}}_{\mathbf{m}}^{\mathbf{I},\mathbf{86/87}} = \hat{\mathbf{M}}_{\mathbf{s}}^{\mathbf{I}}$, where $x = 85$ ($x = 86/87$) in $\hat{\mathbf{M}}_{\mathbf{i}}^{\mathbf{I},\mathbf{x}}$ refers to the 1985 sample (1986/87 sample). In the 1986/87 sample, where we can observe employer/activity changers directly, we find that in 2.31% of (weighted) observations workers changed employers and in 4.65% workers report an activity change within their employers. This implies that more than 93% of the 1985 sample should be made up of employer/activity stayers.

A.2 Results and Discussion

Table 3 shows the occupational transition matrix $\hat{\mathbf{T}}_{\mathbf{s}}^{\mathbf{I}}$ for true occupational stayers derived from $\hat{\mathbf{M}}_{\mathbf{s}}^{\mathbf{I}}$ using the 1985 sample based on the major occupations of the 1990 SOC. This matrix implies that

tional distributions. This exercise yields minimal effects on our statistics. For example, the observed occupational mobility changes by 0.01 percentage point at most. The reason for this small change is because there is the large proportion of the implied true stayers in the sample, which means that $\hat{\mathbf{T}}_{\mathbf{s}}^{\mathbf{I}}$ is not very sensitive to proportional changes in $\hat{\mathbf{M}}_{\mathbf{m}}^{\mathbf{I}}$.

18.46% of true occupational stayers get classified as occupational movers.[34] Although not shown here, a similar conclusion arises when we calculate the same matrix using the 2000 SOC.

Table 10: (Spurious) occupational transition matrix of true occupational stayers, SOC 1990, $\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}}$, (%)

OCCUPATIONS	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) Managing Occ.	**75.7**	3.7	1.0	5.1	8.8	1.5	0.1	0.7	0.7	1.5	0.5	0.4	0.3
(2) Professional Spec.	3.3	**87.9**	2.9	0.4	2.2	1.5	0.1	0.5	0.2	0.6	0.2	0.1	0.1
(3) Technicians	3.1	10.1	**69.5**	0.5	5.9	3.8	0.1	2.7	0.6	1.3	2.1	0.2	0.3
(4) Sales Occ.	6.5	0.5	0.2	**83.3**	3.8	1.5	0.1	0.7	0.2	0.5	0.2	0.8	1.8
(5) Admin. Support	5.8	1.6	1.2	2.0	**84.5**	1.0	0.1	0.3	0.1	0.5	0.7	0.6	1.6
(6) Services	1.7	1.8	1.3	1.3	1.8	**87.2**	0.2	1.3	0.7	0.5	0.6	0.6	0.9
(7) Farm/Fish/Logging	1.3	1.7	0.5	1.3	1.1	2.0	**84.7**	0.6	0.6	0.2	1.0	3.1	2.0
(8) Mechanics	1.8	1.3	2.1	1.3	1.1	2.9	0.1	**79.2**	2.1	2.6	3.0	0.8	1.6
(9) Construction	2.4	0.7	0.6	0.5	0.4	2.2	0.2	3.0	**77.7**	1.7	1.9	1.2	7.6
(10) Precision Prod.	4.5	2.0	1.2	1.1	2.4	1.4	0.0	3.2	1.5	**69.1**	11.1	0.2	2.3
(11) Mach. Operators	0.7	0.3	0.9	0.2	1.4	0.7	0.1	1.6	0.7	4.8	**84.1**	0.7	3.8
(12) Transport	1.1	0.4	0.2	1.6	2.2	1.4	0.7	0.8	0.9	0.2	1.5	**85.1**	3.9
(13) Laborers	1.0	0.4	0.3	4.6	8.2	2.7	0.6	2.1	7.3	2.5	9.7	5.0	**55.7**

A.2.1 The estimate of Γ

Following Proposition *A.1* we can then recover the garbling matrix $\mathbf{\Gamma}$ by using $\hat{\mathbf{T}}^{\mathbf{I}}_{\mathbf{s}}$ and equation (7). Table 4 shows the estimated $\hat{\mathbf{\Gamma}}$ based on the major occupations of the 1990 SOC, while Table 12 shows the estimated $\hat{\mathbf{\Gamma}}$ based on the major occupations of the 2000 SOC. These estimates imply that on average the incorrect occupational code is assigned in around 10% of the cases. Since a spurious transition is likely to be created when either the source or destination occupation is miscoded, the probability of observing a spurious transition for a true occupational stayer is nearly twice as large. Our methodology then suggests that coding error is indeed substantial under independent interviewing. Its magnitude is of similar order as found in other studies analysing the extent of errors in occupational coding (see Campanelli et al., 1997, Sullivan, 2009, and Roys and Taber, 2017).

Two additional messages come out of from Tables 4 and 12. (i) Different occupations have very different propensities to be assigned a wrong code. For example, when using the 1990 SOC we find that individuals whose true occupation is "laborers" have a 74% probability of being coded correctly, while individuals whose true occupation is "professional speciality" have a 94% probability of being coded correctly. (ii) Given a true occupation, some coding mistakes are much more likely than others. For example, workers whose true occupation is "laborers" have a much larger probability to be miscoded as "machine operators" (5.7%), "construction" (4.5%) or "admin. support" (4.8%) than as "managers" (0.3%) or "professionals" (0.1%). Our methodology enable us to take these

[34] This value lies within the expected bounds. To construct an upper bound consider the 1985 sample and assume that all observed occupational transitions are spurious. In this case we can expect that at most 19.71% of the observations would be miscoded. To construct a lower bound consider the 1986/87 sample and calculate the number of observations in which an occupational move is reported among the employer/activity changers. These observations account for 2.60% of all observations in the 1986/87 sample. Assuming that the effect of miscoding is to generate a net increase in the number of occupational changes, we can expect that at least 19.71%-2.60%=17.11% of the observations would be miscoded.

Table 11: Estimate of the garbling matrix, SOC 1990, $\hat{\Gamma}$, (%)

OCCUPATIONS	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) Managing Occ.	**86.8**	2.0	0.5	2.8	4.9	0.8	0.1	0.4	0.4	0.8	0.2	0.2	0.1
(2) Professional Spec.	1.8	**93.7**	1.6	0.2	1.1	0.8	0.1	0.2	0.1	0.3	0.1	0.1	0.0
(3) Technicians	1.7	5.6	**83.3**	0.2	3.3	2.1	0.1	1.5	0.3	0.7	1.1	0.1	0.1
(4) Sales Occ.	3.6	0.2	0.1	**91.2**	1.9	0.8	0.1	0.3	0.1	0.2	0.0	0.4	1.0
(5) Admin. Support	3.2	0.8	0.7	1.0	**91.8**	0.5	0.0	0.1	0.0	0.3	0.3	0.3	1.0
(6) Services	0.9	0.9	0.7	0.7	0.9	**93.3**	0.1	0.7	0.4	0.3	0.3	0.3	0.5
(7) Farm/Fish/Logging	0.7	0.9	0.3	0.7	0.5	1.0	**92.0**	0.3	0.3	0.1	0.5	1.7	1.1
(8) Mechanics	1.0	0.6	1.2	0.7	0.5	1.5	0.1	**88.9**	1.2	1.5	1.6	0.4	0.9
(9) Construction	1.3	0.3	0.3	0.2	0.0	1.1	0.1	1.6	**88.0**	0.9	0.8	0.6	4.7
(10) Precision Prod.	2.6	1.0	0.7	0.6	1.2	0.7	0.0	1.8	0.8	**83.0**	6.3	0.1	1.3
(11) Mach. Operators	0.3	0.1	0.5	0.0	0.7	0.4	0.1	0.8	0.3	2.7	**91.5**	0.4	2.3
(12) Transport	0.5	0.2	0.1	0.8	1.1	0.7	0.4	0.4	0.4	0.0	0.7	**92.2**	2.3
(13) Laborers	0.3	0.1	0.1	2.7	4.8	1.5	0.3	1.2	4.5	1.4	5.7	3.0	**74.3**

differences into account by correcting observed occupational flows by source-destination occupation pair. This provides cleaner net mobility estimates, where the identity of the origin and destination occupation matters.

A.2.2 The effect of $\hat{\Gamma}$ on the gross occupational mobility rate

Figure 12 depicts the effect of $\hat{\Gamma}$ when computing *wave-to-wave* occupational mobility rates. For this exercise we augment the 1985/86/87 samples with workers from the 1984 and 1988 panels, which satisfy the same sample restrictions as before. Figure 12 depicts the average wave-to-wave occupational mobility rate obtained during each year. For the year 1986 we only use the observations that cover the February to December period obtained from the 1986 sample. We label these observations "1986s" as they are the ones we use in our original sample to estimate $\hat{\Gamma}$. In the case of the year 1987, we present the average occupational mobility rate obtained for the January to April period (labelled "1987s") separately from the average occupational mobility rate obtained from the remaining months (labelled "1987r"). The two vertical lines mark the time period in which dependent and independent interviewing overlap.

The short-dashed line with hollow circular markers depicts the observed occupational mobility rates obtained from pooling together employer/activity stayers and changers using the 1984/85 samples, where all respondents were subject to independent interviewing. This pooled sample yields occupational mobility rates that lie between 19.6%-20% and average 19.7% between the vertical lines. Under independent interviewing, $\mathbf{M_m}$ is garbled both at the source and destination occupations and hence is observed as $\mathbf{\Gamma' M_m \Gamma}$. Pre- and post-multiplying the latter by the respective inverses $(\mathbf{\Gamma'})^{-1}, \mathbf{\Gamma}^{-1}$ recovers $(\mathbf{\Gamma'})^{-1}\mathbf{\Gamma' M_m \Gamma \Gamma}^{-1} = \mathbf{M_m}$. Applying this procedure to the 1984/85 samples yields the long-dashed line with hollow triangular markers. The result is a drop in the occupational mobility rate to about 1.6%.

Next consider the 1986/88 samples. Here independent interviewing is only applied to employer/activity changers. Therefore, the observed overall occupational mobility rate for these samples is based on

Table 12: Estimate of the garbling matrix, SOC 2000, $\hat{\mathbf{\Gamma}}$, (%)

OCCUPATIONS	(11)	(13)	(15)	(17)	(19)	(21)	(23)	(25)	(27)	(29)	(31)	(33)	(35)	(37)	(39)	(41)	(43)	(45)	(47)	(49)	(51)	(53)
(11) Management Occ.	**84.2**	2.3	0.3	0.8	0.3	0.3	0.1	0.3	0.1	0.3	0.1	0.1	0.4	0.1	0.4	3.3	3.5	0.5	0.7	0.5	1.2	0.3
(13) Business & Finance Oper.	4.7	**82.8**	0.9	0.5	0.2	0.2	0.2	0.3	0.0	0.0	0.0	0.1	0.0	0.0	0.0	1.6	7.6	0.0	0.1	0.1	0.5	0.1
(15) Computer & Math. Occ.	2.6	3.7	**85.7**	1.6	0.4	0.0	0.0	0.4	0.5	0.0	0.1	0.1	0.0	0.0	0.0	0.0	4.4	0.0	0.1	0.7	0.2	0.2
(17) Architect & Eng. Occ.	2.2	0.7	0.5	**87.0**	1.3	0.0	0.0	0.0	1.2	0.3	0.0	0.0	0.0	0.1	0.0	0.1	1.1	0.0	0.9	2.4	2.2	0.1
(19) Life, Phys, and Soc. Sci. Occ.	2.4	0.6	0.4	3.3	**82.9**	0.0	0.2	1.1	0.4	2.2	0.5	0.0	0.0	0.5	0.0	0.3	1.7	0.7	0.5	1.1	1.0	0.3
(21) Comm & Soc. Service Occ.	2.8	0.7	0.0	0.0	0.0	**89.8**	0.0	1.0	0.4	0.5	0.7	0.3	0.2	0.1	0.4	0.0	2.6	0.0	0.2	0.1	0.1	0.3
(23) Legal	1.3	1.6	0.0	0.0	0.5	0.0	**93.0**	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.1	0.1	3.5	0.0	0.0	0.0	0.0	0.0
(25) Educ., Training & Library	0.5	0.3	0.1	0.0	0.3	0.2	0.0	**97.3**	0.1	0.2	0.1	0.0	0.0	0.0	0.3	0.1	0.6	0.0	0.0	0.0	0.0	0.0
(27) Arts, Dsgn, Ent., Sports & Media	1.0	0.2	0.6	3.8	0.5	0.4	0.0	0.3	**89.2**	0.0	0.0	0.0	0.0	0.0	0.0	0.7	1.3	0.0	0.2	0.2	1.7	0.4
(29) Healthcare Pract. & Tech. Occ.	0.7	0.0	0.0	0.2	0.7	0.1	0.0	0.3	0.0	**92.7**	3.3	0.0	0.1	0.1	0.1	0.1	1.0	0.0	0.0	0.1	0.2	0.1
(31) Healthcare Support	0.5	0.0	0.0	0.0	0.3	0.5	0.0	0.2	0.0	6.8	**88.5**	0.0	0.6	0.4	0.3	0.0	1.6	0.0	0.0	0.1	0.3	0.0
(33) Protective Service	0.7	0.3	0.1	0.0	0.0	0.2	0.2	0.1	0.1	0.1	0.0	**95.0**	0.5	0.1	0.2	0.1	1.8	0.3	0.0	0.1	0.4	0.5
(35) Food Prep/Serving & Rel.	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	0.2	**95.3**	0.5	0.0	1.5	0.4	0.0	0.0	0.0	0.4	0.2
(37) Building/grounds Clean. & Maint.	0.3	0.0	0.0	0.1	0.2	0.1	0.0	0.0	0.0	0.1	0.2	0.0	0.6	**90.3**	0.1	0.4	0.2	0.3	1.7	2.5	1.1	1.8
(39) Personal Care & Service Occ.	3.3	0.0	0.0	0.0	0.0	0.4	0.1	1.6	0.0	0.5	0.6	0.4	0.0	0.3	**91.4**	0.9	0.6	0.0	0.0	0.1	0.0	0.2
(41) Sales & Rel. Occ.	3.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.5	0.1	0.1	**91.7**	1.7	0.1	0.1	0.4	0.3	1.3
(43) Office & Admin. Support	1.7	1.8	0.3	0.2	0.1	0.2	0.1	0.2	0.1	0.2	0.2	0.2	0.1	0.0	0.0	0.9	**91.6**	0.0	0.0	0.1	0.6	1.3
(45) Farm, Fish. & Forestry	4.5	0.2	0.0	0.0	1.0	0.1	0.0	0.1	0.0	0.0	0.0	0.5	0.1	0.9	0.1	0.7	0.3	**87.5**	0.1	0.3	0.8	3.4
(47) Construction & Extraction	1.3	0.1	0.0	0.6	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.2	0.1	0.0	**90.8**	1.6	1.9	2.1
(49) Install., Maint. & Repair Occ.	0.8	0.1	0.1	1.4	0.2	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	1.5	0.0	0.6	0.4	0.1	1.4	**89.0**	3.0	1.1
(51) Production Occ.	0.8	0.2	0.0	0.5	0.1	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.1	0.3	0.0	0.2	0.8	0.0	0.6	1.1	**93.1**	1.9
(53) Transportation & Mater. Moving	0.3	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.1	0.1	0.7	0.0	1.5	2.9	0.4	1.2	0.7	3.4	**88.2**

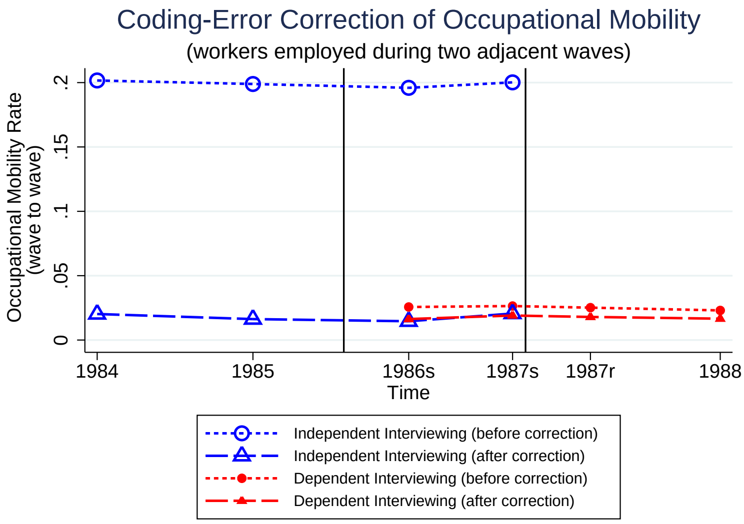


Figure 12: Correcting occupational mobility rates for workers employed in two subsequent waves

$\mathbf{M}_\mathbf{s}^{\mathbf{D},86/87} + \mathbf{M}_\mathbf{m}^{\mathbf{I},86/87}$. The matrix $\mathbf{M}_\mathbf{s}^{\mathbf{D},86/87}$ does not contribute to any occupational transitions while matrix $\mathbf{M}_\mathbf{m}^{\mathbf{I},86/87}$ does. This difference implies a much lower observed aggregate mobility rate. The latter is depicted by the short-dashed line with filled circular markers in Figure 12 and averages 2.6%. Correcting the occupational flows of the 1986-88 samples using $\hat{\mathbf{M}}_\mathbf{s}^{\mathbf{D},86/87} + (\hat{\mathbf{\Gamma}}')^{-1}\hat{\mathbf{M}}_\mathbf{m}^{\mathbf{I},86/87}\hat{\mathbf{\Gamma}}^{-1}$ yields the series depicted by the long-dashed line with solid triangular markers. The result is a drop in the occupational mobility rate to about 1.7%, which is very close to the $\mathbf{\Gamma}$-corrected occupational mobility rate from the 1984/85 samples. Indeed, in Figure 12 the blue (independent-interviewing) and red (dependent-interviewing) long-dashed lines of the $\mathbf{\Gamma}$-corrected measures nearly coincide between the two vertical lines.[35]

The fact that we obtain very similar corrected mobility rates after using the same underlying $\mathbf{\Gamma}$-correction matrix in two different survey designs, suggests that our methodology captures the extent of coding error quite well. Our methodology also seems to work well in other dimensions. We find that the occupational mobility rates in the years before 1986 are adjusted downwards to numbers that are very similar to the numbers obtain during the 1986s-1987s window. For example, the $\mathbf{\Gamma}$-corrected occupational mobility in 1985 is 1.63%, while during the 1986s-1987s window we obtain 1.62%. For the years after the 1986s-1987s window, we find that the $\mathbf{\Gamma}$-corrected occupational mobility rate is within 0.02% of the one obtained during this window. Further, Figure 12 shows that any changes in the level of the $\mathbf{\Gamma}$-corrected occupational mobility rate series appear to track changes in the uncorrected series. This also suggests that our correction method does not seem to introduce additional randomness into the occupational mobility process.

In the main text and the Supplementary Appendix, we apply our $\mathbf{\Gamma}$-correction method to those who changed employers with an intervening spell of unemployment or non-employment. We then compute the host of statistics shown in the main text and Supplementary Appendix. We find that the corrected statistics present a very similar picture as the raw statistics. For example, in the raw data

[35]Note that in the estimation of $\mathbf{\Gamma}$ we used $\hat{\mathbf{M}}_\mathbf{m}^{\mathbf{I},86/87}$, but we did not impose the additional restriction on $(\hat{\mathbf{\Gamma}}')^{-1}\hat{\mathbf{M}}_\mathbf{m}^{\mathbf{I},86/87}\hat{\mathbf{\Gamma}}^{-1}$ to equal $(\hat{\mathbf{\Gamma}}')^{-1}\hat{\mathbf{M}}^{\mathbf{I},85}\hat{\mathbf{\Gamma}}^{-1}$.

we compute an occupational mobility rate at re-employment of 53.1% based on the 2000 SOC. After applying our $\mathbf{\Gamma}$-correction, we obtain an occupational mobility rate at re-employment of 44.4% (see Table 1 in the main text).

A.2.3 The differential impact of coding error on employer/activity stayers and movers

Note that our $\mathbf{\Gamma}$-corrected occupational mobility rate for those workers who changed employers through a spell of unemployment is 16.4% lower than the raw one (44.4% vs 53.1%, 2000 SOC). This adjustment is much smaller in relative terms than the one suggested by Kambourov and Manovskii (2008). They argue that on average around 50% of *all* year-to-year observed occupational mobility in the raw PSID data is due to coding error. Indeed when constructing the year-to-year occupational mobility rate for our pooled sample of employer/activity stayers and movers in the SIPP, we find that only 39.3% of observed occupational changes are genuine (an uncorrected rate of 26.7% vs a $\mathbf{\Gamma}$-corrected rate of 10.5%).

These findings are not mutually inconsistent. The key to this difference is that the *relative* importance of coding error varies greatly with the true propensity of an occupational change among employer stayers and among employer changers. True occupational changes are more likely to be accompanied by changes in employers (for examples of this argument see Hill, 1994, Moscarini and Thomsson, 2007, and Kambourov and Manovskii, 2009) and, vice versa, employer changes are more likely to be accompanied by occupational changes. As such, the *relative* adjustment that Kambourov and Manovskii (2008) find does not automatically carry over to different subsets of workers or flows measured at a different frequency.

As an example, consider an individual who is observed moving from "managers" to "laborers". Suppose "managers" was this individual's true source occupation, but the observed destination occupation was the result of coding error. If this individual is a true occupational stayer, the observed transition will wrongly tell us that the individual *stopped* being a manager and will generate a false occupational move. Instead, if this individual is a true occupational mover, the observed transition, although wrongly coded, will still capture the fact that the individual stopped being a manager and hence will capture a true occupational move. Given that true occupational changes are more likely to occur along side employer changes, there will be more workers among employer changers (relative to employer stayers) whose categorization as an observed occupational mover will not change after using the $\mathbf{\Gamma}$-correction. This implies that the measured occupational mobility rate of employer changers would have a relative smaller adjustment than the measured occupational mobility rate of employer stayers. Kambourov and Manovskii (2008) pooled together employer changers and stayers. Since the latter group represent the vast majority of workers in their sample (as well as in our sample), the relative adjustment proposed by these authors is naturally much larger.

Consider the following iterative back-of-the-envelope approximation to understand why there must be a larger proportion of true occupational movers among those who changed employers than among those who did not change employers. Recall that the $\mathbf{\Gamma}$-correction method implies that true occupational stayers will be coded as movers in about 20% of the times. This happens irrespectively of whether the worker changed employers or not. If we were to suppose that all of the unemployed who regain employment were true occupational stayers, the difference between their observed mobility rate (53%) and coding error (20%) would immediately imply that among the unemployed there must be true occupational movers and these true movers would represent at least 33% of the unemployed. This result then shrinks the population of occupational stayers (the "population at risk") among the unemployed to at most 67%, which (proceeding iteratively) implies the maximum extent of spurious flows produced under the same miscoding propensity is reduced to $0.2 \times 67\%$. In turn, the latter implies an updated lower bound on the percentage of true occupational movers among the unemployed of 39.6%. Proceeding iteratively, one arrives to a lower bound on the percentage of true occupational movers among the unemployed of 41.25%, which is close to the gross mobility we obtained after applying the $\mathbf{\Gamma}$-correction. A similar procedure but applied to employer stayers shows a much lower 'true' mobility rate among this group.

Table 13: Inferred Coding Error Probabilities and Observed vs. Underlying Occupational Mobility

Classification	$\mathbb{P}(\tilde{M}\|S)$	$\mathbb{P}(\tilde{M}\|U)$	$\mathbb{P}(M\|U)$	$\frac{\mathbb{P}(M\|U)}{\mathbb{P}(\tilde{M}\|U)} - 1$	$\mathbb{P}(\tilde{o} \neq o\|U)$
2000 SOC (22 cat)	0.178	0.531	0.444	-0.164	0.095
1990 SOC (13 cat)	0.197	0.507	0.401	-0.209	0.105
1990 SOC (6 cat)	0.148	0.402	0.317	-0.213	0.077
NR/R Cognitive, NR/R Manual (4 cat)	0.110	0.332	0.263	-0.208	0.058
Cognitve, R Manual / NR Manual (3 cat)	0.083	0.273	0.218	-0.199	0.043
Major industry groups (15 cat)	0.101	0.523	0.477	-0.088	0.055

Sample: unemployed between 1983-2013, in 1984-2008 SIPP panels, subject to conditions explained in data construction appendix (most importantly: unimputed occupations (resp. industries), with restrictions to avoid right and left censoring issues.) $\mathbb{P}(\tilde{M}|S)$: probability that the wrong code is assigned to a true stayer; $\mathbb{P}(\tilde{o} \neq o|U)$: probability that the wrong code is assigned to an unemployed worker; $\mathbb{P}(\tilde{M})$: observed occupational mobility among the unemployed; $\mathbb{P}(M)$: inferred underlying true mobility (proportion of unemployed). *NR/R* refers to routine vs. non-routine. Further details on the classifications are explained in the data construction appendix.

Table 13 shows that the differential impact of coding error on employer changers and stayers is present when considering several alternative occupational classifications as well as mobility across industries. In all these cases we use wave-to-wave mobility rates. The first column, $\mathbb{P}(\tilde{M}|S)$, presents the probability that a true employer/activity stayer in the 1985 sample is assigned the wrong occupational code and hence is observed as an occupational mover $\tilde{M}$. This probability is 17.8% when using the 2000 SOC. This implies that under independent interviewing we will observe an occupational mobility rate of 17.8% among the employer/activity stayers. This rate increases slightly when using the 1990 SOC, 19.7%, and remains high even when we aggregate occupations into six categories, 14.8%.[36] Aggregating occupations into four tasked based categories (routine vs. non-routine and manual vs. cognitive) only brings down the observed mobility rate of true occupational stayers to 11%. As many other studies, we also find the probability that a true stayer is observed as a mover is lower when considering industries instead of occupations.

The second and third columns show the observed, $\mathbb{P}(\tilde{M}|U)$, and $\mathbf{\Gamma}$-corrected, $\mathbb{P}(M|U)$, occupational mobility rate of those workers who changed employers through unemployment. These are obtained using the probability that an unemployed worker in the 1984-2008 SIPP panels is assigned

[36]The six groups are: (1) managers/professional speciality; (2) tech support/admin support/sales; (3) services; (4) farm/forest/fisheries; (5) precision production/craft/repair; and (6) operators, fabricators and laborers. These correspond to the summary occupational group of the 1990 SOC.

the wrong occupational code at re-employment, $\mathbb{P}(\tilde{o} \neq o|U)$. We observe that across all classifications the (relative) difference between the observed and the $\mathbf{\Gamma}$-corrected occupational mobility rates of the unemployed, $\mathbb{P}(\tilde{M}|U) - \mathbb{P}(M|U)$, is about half the size of $\mathbb{P}(\tilde{M}|S)$. The fourth column shows this difference in relative terms. It is clear that across all classifications the same coding error generates a much larger difference (in absolute and relative terms) between the observed and $\mathbf{\Gamma}$-corrected occupational mobility rates of employer stayers than among employer changers.

A.3 Measuring coding error in the PSID

We now broaden the above analysis and use probabilistic models based on the PSID (as in Kambourov and Manovskii, 2008) to assess the impact of coding errors on the probability of an occupational change. The advantage of the $\mathbf{\Gamma}$-correction method is that it captures all sources of error in assigning occupation codes. It delivers an identification procedure that is not subject to the issue highlighted in Corollary A.1 and that recovers the extent to which coding errors arise at the level of each occupation. Further, the SIPP provides a large sample size in which we can apply our correction method. The advantage of the PSID is that retrospective coding affected directly the way occupational transitions among employer changers were measured. We exploit this feature and assess the impact of coding error on employer movers and employer stayers separately and compare the results with the ones obtained from the SIPP using our $\mathbf{\Gamma}$-correction method. We find a very consistent picture across the two data sets.

To assess the impact of retrospective coding in reducing coding errors, we use the PSID retrospective occupation-industry supplementary data files, which contain the re-coding the PSID staff performed on the occupational mobility records obtained during the 1968-1980 period. Since the 1981-1997 records were not re-coded and collected under independent interviewing, the earlier period can be used to construct "clean" occupational mobility rates and to analyse the effect of measurement error at the coding stage. In constructing our sample we closely follow Kambourov and Manovskii (2008, 2009). The details of this sample are described in the Supplementary Appendix G.1.

A.3.1 Gross occupational mobility rates

As in Kambourov and Manovskii (2008) we define the *overall* occupational mobility rate as the fraction of employed workers whose occupational code differs between years t and $t+1$ divided by the number of workers who were employed in year t. As these authors we also consider those workers who were employed at the time of the interview in year $t-1$, unemployed in t and employed at the time of the interview in year $t+1$. Further, we define the *within-employer* occupational mobility rate as the fraction of workers employed who did not change employers but exhibit a different occupational code between years t and $t+1$, divided by the number of employed workers who did not change employers between years t and $t+1$. Similarly, we define the *across-employer* occupational mobility rate be the fraction of employed workers who's occupational code differs between years t and $t+1$ and reported an employer change between these years, divided by the number of employed workers in year t who have reported an employer change between years t and $t+1$. To identify employer changes we follow the procedure detailed in Kambourov and Manovskii (2009), Appendix A1.

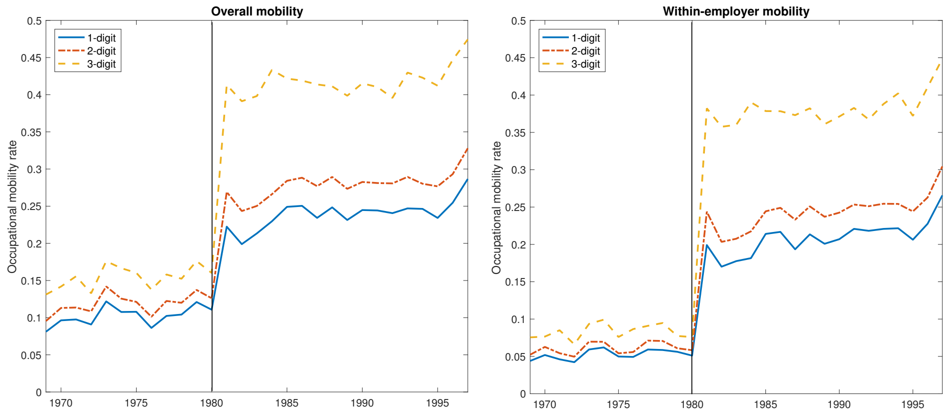


Figure 13: Overall and within-employer occupational mobility rates

The left panel of Figure 13 depicts the yearly overall occupational mobility rate at a one-, two- and three-digit level of aggregation, using the 1970 SOC. When retrospective re-coding was used, the overall occupational mobility rate experienced a large downward shift, ranging between 10 to 25 percentage points, depending on the level of aggregation of the occupational codes. These drops suggest that only between 38% to 45% of all occupational moves are genuine. This is very similar to the conclusion reached by Kambourov and Manovskii (2008).

The right panel of Figure 13 shows that the *within-employer* occupational mobility rates experienced even stronger drops than the overall ones under retrospective coding. In contrast, the left panel of Figure 14 shows that the impact of coding error in the *across-employer* occupational mobility rates is much more moderate and hardly visible when aggregating occupations at a one-digit level. These results then suggest that the impact of coding error on the overall mobility rates mainly arises from those workers who did not change employers, where employer stayers account on average for 87.1% of all employed workers in a given year, while those who changed employers account for the remainder 12.9%. We find a similar conclusion based on the SIPP data.

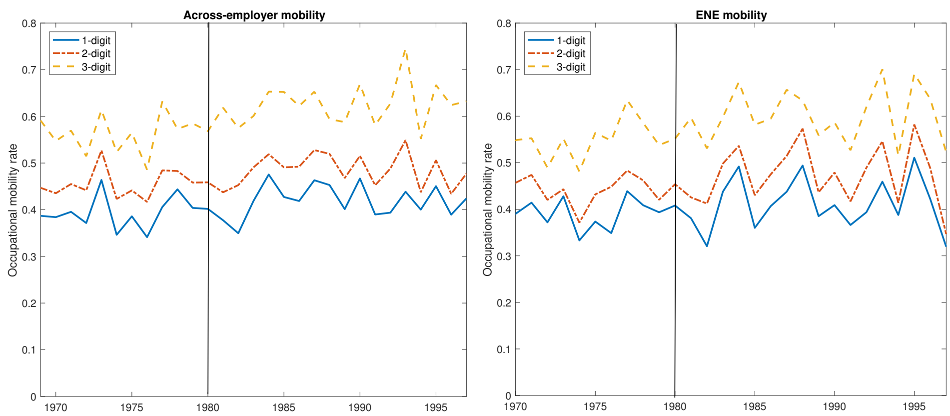


Figure 14: Across-employer occupational mobility rates

Next consider the effects of coding error on the occupational mobility rate of only those workers who changed employers through a spell of non-employment (ENE). To construct the ENE occupational mobility rate we consider (i) those workers who were employed at the interview date in year $t-1$, non-employed at the interview date in year t and once again employed at the interview date in year $t+1$; and (ii) those workers employed at the interview dates in years t and $t+1$, but who declared that they experienced an *involuntary* employer change between these two interviews. An involuntary change is defined as those cases were the worker declared a job separation due to "business or plant closing", due to "being laid off or were fired" or their "temporary job ended" (see Supplementary Appendix G.1 for details). We divide these flows by the number of workers who changed employers through a spell of non-employment during the corresponding years. The right panel of Figure 14 shows that coder error once again seems to have a small effect on the occupational mobility rates of these workers.

A.3.2 Probabilistic models

The visual impressions given by the above figures on the effects of coding error are confirmed when estimating the effects of retrospective coding using a probit or a linear probability model. In these regressions the dependent variable takes the value of one if the worker changed occupation and zero otherwise. We include the indicator variable "break" which takes the value of one during the years in which the PSID used retrospective coding. In addition we control for age, education, full or part-time work, occupation of origin, region of residence, aggregate and regional unemployment rates, a quadratic time trend and number of children.[37]

Table 14 shows the marginal effects for the probit regressions.[38] It shows that retrospective coding had a large and significant effect on reducing the probability of changing occupations when all workers were included in the sample. Furthermore, the values of the marginal effects of the "break" indicator are very close to the amount by which the overall occupational mobility series shifted when retrospective coding was used, as depicted in Figure 13.

Our estimates also show that the effect of retrospective coding is much more moderate when we condition the sample on workers who changed employers and when we consider the ENE sample.[39]

A.3.3 A comparison of coding errors across the PSID and SIPP

The point estimates obtained in Table 14 suggest that the probability of an occupational change for those who changed employers and for those who changed employer through non-employment spells,

[37] As in Kambourov and Manovskii (2008), the education indicator variable takes the value of one when the worker has more than 12 years of education and zero otherwise. This is to avoid small sample problems if we were to divide educational attainment in more categories. The regional unemployment rates are computed using US states unemployment rates.

[38] These estimates are obtained using the personal weights provided by each survey, but similar results are obtained when using the unweighted data. We also obtained very similar results when using the linear probability model on weighted and unweighted data and when using robust standard errors and clustering standard errors at a yearly level.

[39] We do not include the within-employer occupational mobility in Table 14 because, as suggested by the graphical analysis, the results are very similar to the ones obtained with the full sample.

Table 14: The effect of measurement error on the PSID (probit marginal effects)

	All workers			Across employer			ENE		
	1-digit	2-digits	3-digits	1-digit	2-digits	3-digits	1-digit	2-digits	3-digits
Unemp rate	-0.002	-0.004	0.001	-0.026**	-0.025**	-0.021**	-0.058***	-0.057***	-0.032*
Reg unemp rate	-0.003	-0.003	-0.003	0.015**	0.014*	0.013*	0.046***	0.039***	0.022*
Age	-0.008***	-0.008***	-0.011***	-0.007	-0.008	-0.007	-0.025*	-0.029**	-0.022*
Age squared	0.6 e-4***	0.6 e-4**	0.9 e-4***	0.4 e-4	0.3 e-4	0.2 e-4	0.26 e-4	0.29 e-4*	0.19 e-3
Education	0.015***	0.016***	0.007	0.026*	0.026*	0.030*	0.024	0.034	0.069**
Break	-0.133***	-0.165***	-0.260***	-0.040	-0.069**	-0.065**	-0.029	-0.077	-0.065
Full-time	0.034***	0.017	-0.001	-0.019	-0.055	-0.101***	0.051	0.043	-0.065
N Obs	39,047	39,047	38,841	4,962	4,935	4,656	1,792	1,782	1,576
R^2	0.010	0.119	0.184	0.040	0.067	0.134	0.064	0.081	0.175

Levels of significance: $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

should be lowered on average by 3 percentage points at a one-digit level, 8 percentage points at a two-digit level and 7 percentage points at a three-digit level to capture the effect of coding error. To compare these estimates to the ones obtained from the SIPP, first note that retrospective coding is done using the same descriptions of the "kind of work" individuals gave in past interviews and hence captures coding errors introduced at the coding stage (see Sullivan, 2010). In addition to this coder error, our $\mathbf{\Gamma}$-correction method also takes into account that the source of code disagreement can originate in different descriptions of the same work (respondent error). Hence we would expect a higher correction when using the $\mathbf{\Gamma}$-correction than when using retrospective coding. Taking this feature into account and noting that coder error is expected to be the most important source error (see Mathiowetz, 1992), the PSID estimates compare very well with the adjustments implied by the $\mathbf{\Gamma}$-correction method.

In particular, when aggregating occupations into major categories (2000 SOC or 1990 SOC) and using the $\mathbf{\Gamma}$-correction, the corrected average occupational mobility rate for the non-employed was approximately 11 percentage points lower than the one obtained using the raw SIPP data. Noting that the occupational aggregations used in the SIPP and the two-digit aggregation used in the PSID lead to very similar ENE occupational mobility rates, the 11 percentage point adjustment obtained from the SIPP is thus close to the 8 percentage points suggested by Table 14.[40] The difference between the adjustments obtained from the SIPP and the PSID (3 percentage points) then provides a rough estimate of the impact of the respondent error. This estimate then implies that the importance of coder error is about 2.6 times larger than the importance of the respondent error. This is remarkably consistent with Mathiowetz (1992), who shows that the importance of coder error is two times larger than the importance of the respondent error when aggregating codes at a one-digit level and five times larger than the importance of the respondent error when aggregating codes at a three-digit level.

Furthermore, both the SIPP and the PSID data sets strongly suggest that the percentage reduction of occupational mobility due to coding error varies substantially between employer stayers and employer movers. In the case of employer stayers, a large percentage of transitions are implied to be

[40]For the 1985-1995 period, during which the PSID and SIPP overlap, the average "year-to-year" occupational mobility rate of the non-employed in the PSID and the SIPP was both around 53.1% and 47.6% when using the major occupational categories of the 1990 SOC.

spurious. In the PSID we find that at a two-digit level 45% of yearly transitions are spurious, while in the SIPP we find that 40% of the yearly transitions of employer stayers are spurious. In the case of employer movers, occupational mobility is reduced by about 10 percentage points, but high occupational mobility remains (around 40% comparing before and after an employer changers), after applying retrospective coding or after using the $\mathbf{\Gamma}$-correction. As discussed earlier, this difference arises as among employer stayers the proportion of true occupational stayers is high and coding errors translate into a large amount of spurious mobility. Among those who changed employers through non-employment there is a much smaller proportion of true occupational stayers and hence the "population at risk" to be assigned a spurious occupational change is smaller.

B Theory

This appendix complements Section 3 of the paper. The first part investigates, using a simplified version of the model, (i) the conditions under which rest unemployment arises and (ii) the cyclical properties of workers' job separation and occupational mobility decisions. The second part presents the equations describing worker flows in a BRE. The third part provides the definition of a BRE, the proof of Proposition 2 (in the main text), the proof of existence of the separation and reallocation cutoffs. However, to safe space the proofs of the results discussed in the first part of the appendix can be access by following the coloured link: Appendix B and through the authors' personal websites.

B.1 Model Implications and Comparative Statics

We start by exploring the main implications of our theory: the occurrence of rest unemployment and the cyclical properties of workers' decisions to search across occupations and to separate from jobs. To keep the intuition as clear as possible we study a simplified version of our model without occupational human capital accumulation, setting $x_h = 1$ for all h, and without occupational-wide productivity shocks, setting $p_o = 1$ for all o. These restrictions imply that all occupations are identical, worker mobility across occupations is fully undirected and purely driven by occupation-worker idiosyncratic shocks such that gross mobility equals excess mobility. Further, within an occupation workers differ only in their z-productivities and labor market segmentation is done along this dimension. Agents' value functions are still given by the Bellman equations described in Section 3.2 of the main text, but now with state space (z, A) instead of $(z, x, o, A, \mathcal{P}_O)$. The net value of searching across occupations then simplifies to

$$R(A) = -c + \int_{\underline{z}}^{\overline{z}} W^U(A, \tilde{z}) dF(\tilde{z}). \tag{8}$$

In Carrillo-Tudela and Visschers (2013) we show that in this setting and assuming $F(z'|z) < F(z'|\tilde{z})$ for all z, z' if $z > \tilde{z}$, the value functions $W^U(A, z)$, $W^E(A, z)$, $J(A, z)$ and $M(A, z)$, exist, are unique and increase in z. This implies that $\theta(A, z)$ also exists, is unique and increases with z. Further we show that if $\delta + \lambda(\theta(A, z)) < 1$ for all A, z, in equilibrium there exists a unique cutoff function z^s that depends only on A, such that $d(A, z) = \sigma(A, z) = 1$ if and only if $z < z^s(A)$, and

$d(A,z)=\sigma(A,z)=0$ otherwise. Since $R(A)$ is constant in z, $W^U(A,z)$ is increasing in z, and by the existence of a unique z^s so is $\max\{\lambda(\theta(A,z^r(A)))(W^E(A,z^r(A))-W^U(A,z^r(A))),0\}$, there also exists a reallocation cutoff function $z^r(A)$ such that workers decide to search across occupations if and only if $z<z^r(A)$ for every A, where $z^r(A)$ satisfies

$$R(A)=W^U(A,z^r(A))+\max\{\lambda(\theta(A,z^r(A)))(W^E(A,z^r(A))-W^U(A,z^r(A))),0\}. \tag{9}$$

Using this simplified framework we first study the relative positions of the job separation and reallocation cut-off functions and hence gain insights on the conditions under which rest unemployment arises. We then study the slopes of these cut-off functions and gain insights into the cyclicality of separations and (excess) occupational mobility in our model. Using our calibrated model we have verified that the same properties as derived below apply to the more general setup considered there.

B.1.1 The Occurrence of Rest Unemployment

We first analyse how the value of waiting in unemployment and in employment changes with c, b and the persistence of the z-productivity process, and how these changes determine the relative position of z^r and z^s. The simplest setting that captures a motive for waiting is one in which the z-productivity is redrawn randomly with probability $0<(1-\gamma)<1$ each period from cdf $F(z)$ and A is held constant. Time-variation in z is essential here because a worker can decide to stay unemployed in his occupation, even though there are no jobs currently available for him, when there is a high enough probability that his z-productivity will become sufficiently high in the future. All other features of the model remain the same, with the exception that we do not consider human capital accumulation or occupational-wide productivity differences. In Section 3.2 of this appendix we formulate the value functions for this stationary environment and provide the proofs of Lemmas 1 and 2, below.

In this stationary setting, the expected value of an unemployed worker with productivity z, measured at the production stage, is given by

$$\begin{aligned} W^U(A,z) &= \gamma\Big(b+\beta\max\Big\{R(A),W^U(A,z)+\max\{\lambda(\theta(A,z))(1-\eta)(M(A,z)-W^U(A,z)),0\}\Big\}\Big) \\ &+ (1-\gamma)\mathbb{E}_z[W^U(A,z)]. \end{aligned} \tag{10}$$

Equation (10) shows that there are two ways in which an unemployed worker with a $z<z^s$ can return to production. *Passively*, he can wait until his z-productivity increases exogenously. Or, *actively*, by paying c and sampling a new z in a different occupation. In the case in which the worker prefers to wait, the inner $\max\{.\}=0$ as the workers is below the separation cutoff and the outer $\max\{.\}=W^U(A,z)=W^U(A,z^s(A))$, where the latter equality follows as in this simplified environment the z-productivity process is assumed to be iid. In the case in which the worker prefers to search across occupations the outer $\max\{.\}=R(A)=W^U(A,z^r(A))$. The difference $W^U(A,z^s(A))-R(A)$, then captures the relative gain of waiting for one period over actively sampling a new z immediately. If $W^U(A,z^s(A))-R(A)\geq 0$, then $z^s\geq z^r$ and there is rest unemployment. If $W^U(A,z^s(A))-R(A)<0$, then $z^r>z^s$, and endogenously separated workers immediately search across occupations.

Changing c, b, or γ will affect the relative gains of waiting, in employment and in unemployment. In the following lemma we derive the direction of the change in $W^U(A,z^s(A))-R(A)$, where we take fully into account the feedback effect of changes in c, b, or γ on the match surplus $M(A,z) - W^U(A,z)$ that arises due to the presence of search frictions as discussed in Section 3.3 in the main text.

Lemma 1. *Changes in c, b or γ imply*

$$\frac{d(W^U(A,z^s(A)) - R(A))}{dc} > 0, \frac{d(W^U(A,z^s(A)) - R(A))}{db} > 0, \frac{d(W^U(A,z^s(A)) - R(A))}{d\gamma} < 0.$$

It is intuitive that raising c directly increases the relative gains of waiting as it make occupational mobility more costly. An increase in c, however, also leads to a larger match surplus because it reduces $W^U(A,z)$, making employed workers less likely to separate and hence reducing rest unemployment. The lemma shows that, overall, the first effect dominates. A rise in b lowers the effective cost of waiting, while at the same time decreasing the match surplus by increasing $W^U(A,z)$, pushing towards more rest unemployment. An increase in γ, decreases the gains of waiting because it decreases the probability of experiencing a z-shock without paying c and hence increases the value of sampling a good z-productivity. In the proof of Lemma 1 we further show that an increase in $W^U(A,z^s(A)) - R(A)$ leads to an increase in $z^s(A) - z^r(A)$. This implies that for a sufficiently large c, b or $1-\gamma$ rest unemployment arises.

Rest Unemployment and Occupational Human Capital Occupational human capital accumulation makes a worker more productive in his current occupation. This implies that workers are willing to stay longer unemployed in their occupations because, for a given z, they can find jobs faster and receive higher wages. At the same time, a higher x makes the employed worker less likely to quit into unemployment, generating a force against rest unemployment. Taken together, however, the first effect dominates as the next result shows.

Lemma 2. *Consider a setting where A is fixed, z redrawn with probability $(1-\gamma)$, and production is given by $y = Axz$. Consider an unexpected, one-time, permanent increase in occupation-specific human capital, x, from $x = 1$. Then*

$$\frac{d(W^U(A,z^s(A,x),x) - R(A))}{dx} > 0.$$

This result implies that the difference $z^s(A,x) - z^r(A,x)$ becomes larger when human capital increases. Thus, more occupational human capital leads to rest unemployment.

B.1.2 The Cyclicality of Occupational Mobility and Job Separation Decisions

In Section 2.5 of the main text we documented that occupational mobility through unemployment (non-employment) is procyclical, while it is well established that job separations into unemployment (non-employment) are countercyclical (see Section 5 in the main text). In the model, the cyclicalities of workers' occupational mobility and job separation decisions are characterised by the slopes of the cutoff functions z^r and z^s with respect to A. As discussed in Section 3.3 of the main text, occupational mobility decisions are procyclical and job separation decisions are countercyclical when $dz^r/dA > 0$ and $dz^s/dA < 0$, respectively. We now explore the conditions under which such slopes arise endogenously in our model.

Occupational mobility decisions We start by investigating the impact of rest unemployment on the slope of z^r using the simplified version of the model (i.e. $x_h = 1$ for all h and $p_o = 1$ for all o). Using (9) and noting that $R(A) - W^U(A, z^r(A)) = 0$, we obtain

$$\frac{dz^r}{dA} = \frac{\int_{z^r}^{\overline{z}} \left(\frac{\partial W^U(A,z)}{\partial A} - \frac{\partial W^U(A,z^r)}{\partial A}\right) dF(z) - \frac{\partial}{\partial A}\Big(\lambda(\theta(A,z^r))(W^E(A,z^r) - W^U(A,z^r))\Big)}{\frac{\partial W^U(A,z^r)}{\partial A} + \frac{\partial}{\partial z^r}\Big(\lambda(\theta(A,z^r))(W^E(A,z^r) - W^U(A,z^r))\Big)}. \quad (11)$$

Since workers who decided to search across occupations must sit out one period unemployed, the term

$$\lambda(\theta(A, z^r))(W^E(A, z^r) - W^U(A, z^r))$$

captures the expected loss associated with the time cost of this decision: by deciding not to search across occupations, the worker could match with vacancies this period. When $z^r > z^s$, $\lambda(\theta(A, z^r))(W^E(A, z^r) - W^U(A, z^r)) > 0$ and is increasing in A and z^r. Therefore, an increase in A in this case, increases the loss associated with the time cost of searching across occupations and decreases dz^r/dA. However, when rest unemployment occurs ($z^s > z^r$), $\lambda(\theta(A, z^r))(W^E(A, z^r) - W^U(A, z^r)) = 0$ and this effect disappears. This follows as during rest unemployment periods workers have a contemporaneous job finding rate of zero and hence by searching across occupations, the worker does not lose out on the possibility of matching this period. In this case, the cyclicality of occupational mobility decisions purely depends on the remaining terms, in particular $\int_{z^r}^{\overline{z}} \left(\frac{\partial W^U(A,z)}{\partial A} - \frac{\partial W^U(A,z^r)}{\partial A}\right) dF(z)$. Thus, the presence of rest unemployment adds a procyclical force to occupational mobility decisions.

Now consider the impact of search frictions on the slope of z^r. We focus on the more general case of $z^r > z^s$, which includes the additional countercyclical force discussed above. To gain analytical tractability we consider the stationary environment used in the previous subsection and set $\gamma = 1$ such that both A and workers' z-productivities are permanent. This allows us to link wages and labor tightness to $y(A, z)$ in closed form. We then analyse the effects of a one-time, unexpected, and permanent change in A on z^r.[41] To isolate the role of search frictions, we compare this case with one without search frictions in which workers (who are currently not changing occupations) can match instantaneously with firms and are paid $y(A, z)$. In both cases we keep in place the same reallocation frictions. Let z_c^r denote the reallocation cutoff in the case without search frictions. The details of both cases, including the corresponding value functions and the proof of the following lemma can be found in Supplementary Appendix H.

Lemma 3. *Consider a one-time, unexpected, permanent increase in A. With search frictions the cyclical response of the decision to search across occupations is given by*

$$\frac{dz^r}{dA} = \frac{\beta \int_{z^r}^{\overline{z}} \Big(\big(\frac{C_s(A,z)}{C_s(A,z^r)}\big) y_A(A,z) - y_A(A,z^r)\Big) dF(z) - (1-\beta) y_A(A,z^r)}{(1 - \beta F(z^r)) y_z(A,z^r)}, \quad (12)$$

[41] This approach follows Shimer (2005), Mortensen and Nagypal (2007), and Hagedorn and Manovskii (2008). Since the equilibrium value and policy functions only depend on A and z, analysing the change in the expected value of unemployment and joint value of the match after a one-time productivity shock is equivalent to compare those values at the steady states associated with each productivity level. This is because in our model the value and policy functions jump immediately to their steady state level, while the distribution of unemployed and employed over occupations takes time to adjust.

where $y_i(A,z) = \partial y(A,z)/\partial i$ *for* $i = A, z$ *and* $C_s(A,z) = \frac{\beta\lambda(\theta(A,z))}{(1-\beta)(1-\beta+\beta\lambda(\theta(A,z)))}$. *Without search frictions the cyclical response is given by*

$$\frac{dz_c^r}{dA} = \frac{\beta \int_{z_c^r}^{\bar{z}} (y_A(A,z) - y_A(A,z_c^r))dF(z) - (1-\beta)y_A(A,z_c^r)}{(1-\beta F(z_c^r))y_z(A,z_c^r)}. \tag{13}$$

The first term in the numerator of (12) and (13) relates to $\int_{z^r}^{\bar{z}} \left(\frac{\partial W^U(A,z)}{\partial A} - \frac{\partial W^U(A,z^r)}{\partial A}\right) dF(z)$ in equation (11), while the second term captures the opportunity cost of the reallocation time. The proof of Lemma 3 shows that $C_s(A,z)/C_s(A,z^r)$ is increasing in z and $\frac{dz^r}{dA} > \frac{dz_c^r}{dA}$ at $z^r = z_c^r$. Thus, the presence of search frictions also adds procyclicality to the decision to search across occupations. Search frictions lead to a steeper reallocation cutoff function because, in this case, an increase in A increases $W^U(A,z)$ through *both* the wage and the job finding probability. In contrast, an increase in A in the frictionless case only affects $W^U(A,z)$ through wages (with a proportionally smaller effect on $w - b$), as workers always find jobs with probability one. The fact that $C_s(A,z)/C_s(A,z^r)$ is increasing in z for $z > z^r$, reflects that the impact of $y(A,z)$ on $W^U(A,z)$ is increasing in z. Indeed, from the proof of Lemma 3 one obtains $\frac{\partial W^U(A,z)}{\partial A} \Big/ \frac{\partial W^U(A,z^r)}{\partial A} = \frac{C_s(A,z)y_A(A,z)}{C_s(A,z^r)y_A(A,z^r)} > \frac{y_A(A,z)}{y_A(A,z^r)}$ for $z > z^r$.

One can get further intuition by considering the planners' problem (for details see Carrillo-Tudela and Visschers, 2013). The envelope condition implies that the planner, at the optimum allocation, does not need to change labor market tightness at each z for a infinitesimal change in A to still obtain the maximum net increase in expected output. At a given z, this means that an increase of dA creates

$$dW^U(A,z) = \frac{\beta\lambda(\theta(A,z))}{(1-\beta)(1-\beta+\beta\lambda(\theta(A,z)))} y_A(A,z)dA = C_s(A,z)y_A(A,z)dA$$

in additional life-time expected discounted output for the planner. Since our economy is constrained efficient, the change in A also creates the same lifetime expected discounted income for an unemployed worker with such z.

In addition, (12) and (13) show that the cyclicality of z^r in either case depends on the production function $y(A,z)$, in particular on the sign of $y_A(A,z) - y_A(A,z^r)$ for $z > z^r$. In the proof of Lemma 3 we show that with search frictions, the decisions to search across occupations will already be procyclical when the production function is modular and z^r is sufficiently close to z^s. This follows because the opportunity cost of searching across occupations becomes smaller as z^r approaches z^s from above. With rest unemployment this opportunity cost is zero. If z^r is substantially above z^s, we will need sufficient complementarities between A and z in the production function to obtain a procyclical z^r. Without search frictions, in contrast, a supermodular production function is only a necessary condition to generate procyclicality in z^r .

Job Separations The main aspect of having endogenous job separation and occupational mobility decisions is that the two can potentially interact. In particular, if $z^r > z^s$, workers separate endogenously to search across occupations and this could lead to z^r and z^s having the same cyclical behavior. For example, in the setting of Lemma 3 were both A and workers' z-productivities are permanent, we show at the end of Section 3.2 of this appendix that dz^s/dA depends directly on dz^r/dA. Namely,

$$\frac{dz^s(A)}{dA} = -\frac{y_A(A,z^s(A))}{y_z(A,z^s(A))} + \frac{\beta\lambda(\theta(A,z^r(A)))}{1-\beta(1-\delta)+\beta\lambda(\theta(A,z^r(A)))} \frac{y_A(A,z^r(A))}{y_z(A,z^s(A))} \left(1 + \frac{y_z(A,z^r(A))}{y_A(A,z^r(A))}\frac{dz^r(A)}{dA}\right).$$

The second term makes explicit the interaction between the decisions to separate from a job and to search across occupations when there is no rest unemployment. It captures the change in the gains of search across occupations, $dR(A)/dA$, and shows that z^r and z^s can have the same cyclicality. When instead $z^s > z^r$, workers endogenously separate into a period of rest unemployment. In this case, $R(A)$ has a smaller impact on the value of unemployment at the moment of separation. This is because searching across occupations would only occur further in the future, and then only if a worker's z-productivity would deteriorate below z^r. Thus, the presence of rest unemployment weakens any feedback of a procyclical z^r onto z^s. Indeed, by setting $\lambda(.) = 0$ in the above expression we get that z^s is always countercyclical.

B.2 Worker Flows

In a BRE the evolution of the distribution $\mathcal{G}$ of workers across labor markets (z, x), occupations o and employment status es is a result of (i) optimal vacancy posting $\theta(.)$, job separation decisions $d(.)$ and occupational mobility decisions $\rho(.)$ and $\mathcal{S}(.)$, all depending on the state vector $\omega = (A, \mathcal{P}_O, z, x, o)$; and (ii) the exogenous retiring probability μ. To obtain the laws of motions of unemployed and employed workers it is then useful to derive the measure of unemployed and employed workers at each stage j within a period, where $j = s, r, m, p$ represent separations, reallocations, search and matching and production as described in the main text. Let $u_t^j(z, x_h, o)$ denote the measure of unemployed workers in labor market (z, x_h) in occupation o at the beginning of stage j in period t. Similarly, let $e_t^j(z, x_h, o)$ denote the measure of employed workers in labor market (z, x_h) in occupations o at the beginning of stage j in period t.

B.2.1 Unemployed workers

Given the initial conditions $(A_0, \mathcal{P}_{O,0}, \mathcal{G}_0^p)$, the measure of unemployed workers characterised by (z, x_h) in occupation o at the beginning of next period's separation stage is

$$\begin{aligned} u_{t+1}^s(z, x_h, o)dz &= (1-\mu)\Bigg[\chi^u(x_h|x_h)\int_{\underline{z}}^{\overline{z}} u_t^p(\tilde{z}, x_h, o)dF(z|\tilde{z})d\tilde{z} + \chi^u(x_h|x_{h+1})\int_{\underline{z}}^{\overline{z}} u_t^p(\tilde{z}, x_{h+1}, o)dF(z|\tilde{z})d\tilde{z} \\ &+ \mu\Bigg[\sum_{\tilde{o}=1}^{O}\sum_{\tilde{h}=1}^{H}\int_{\underline{z}}^{\overline{z}}\Big[u_t^p(\tilde{z}, x_{\tilde{h}}, \tilde{o}) + e_t^p(\tilde{z}, x_{\tilde{h}}, \tilde{o})\Big]d\tilde{z}\Bigg]\psi_o(\mathbf{1}_{h=1})dF(z). \qquad (14 \end{aligned}$$

Conditional on not retiring from the labor market, the terms inside the first squared bracket show the probability that unemployed workers in labor markets $(\tilde{z}, x_h, o)$ and $(\tilde{z}, x_{h+1}, o)$ in the previous period's production stage will be in labor market (z, x_h, o) immediately after the z and x_h shocks occur. The term in the second squared bracket refers to the measure of new workers who entered the economy to replace those who left at the beginning of the period due to the μ-shock. We assume that the population of workers is constant over time, making the inflow equal to the outflow of workers. New workers are assumed to enter unemployed with a $\tilde{z}$ randomly drawn from $F(.)$ and with the lowest human capital level x_1. The above equation considers the inflow who has been assigned productivity z and occupation o, where ψ_o denotes the probability that workers in the inflow are assigned occupation o and $\mathbf{1}_{h=1}$ denotes an indicator function which takes the value of one when the labor market (z, x_h) is associated with x_1 and zero otherwise.

During the separation stage some employed workers will become unemployed. Since by assumption these newly unemployed workers do not participate in the current period's reallocation or search and matching stages, it is convenient to count them at the production stage. This implies that $u^s_{t+1}(z, x_h, o)dz = u^r_{t+1}(z, x_h, o)dz$. Similarly, we will count at the production stage those unemployed workers who arrived from other occupations during the reallocation stage, as they also do not participate in the current period's search and matching stage. This implies that the measure of unemployed workers characterised by (z, x_h) in occupation o at the beginning of the search and matching stage is given by

$$u^m_{t+1}(z, x_h, o)dz = (1 - \rho(A, \mathcal{P}_O, z, x_h, o))u^r_{t+1}(z, x_h, o)dz.$$

Noting that $(1-\lambda(\theta(A, \mathcal{P}_O, z, x_h, o)))u^m_{t+1}(z, x_h, o)dz$ workers remain unemployed after the search and matching stage, the above assumptions on when do we count occupational movers and those who separated from their employers imply that the measure of unemployed workers characterised by (z, x_h) in occupation o during the production stage is given by

$$u^p_{t+1}(z, x_h, o)dz = (1-\lambda(\theta(A, \mathcal{P}_O, z, x_h, o)))u^m_{t+1}(z, x_h, o)dz + d(z, x_h, o, A, \mathcal{P}_O)e^s_{t+1}(z, x_h, o)dz \tag{15}$$

$$+ (\mathbf{1}_{h=1})\Bigg[\sum_{\tilde{o}\neq o}\sum_{\tilde{h}=1}^{H}\Big[\int_{\underline{z}}^{\overline{z}}\rho(\tilde{z}, \tilde{x}_h, \tilde{o}, A, \mathcal{P}_O)\alpha(s_o(\tilde{z}, \tilde{x}_h, \tilde{o}, A, \mathcal{P}_O), \tilde{o})u^r_{t+1}(\tilde{z}, x_{\tilde{h}}, \tilde{o})d\tilde{z}\Big]dF(z)\Bigg].$$

B.2.2 Employed workers

Next we turn to describe the laws of motion for employed workers. Given the initial conditions $(A_0, \mathcal{P}_{O,0}, \mathcal{G}^p_0)$, the measure of employed workers characterised by (z, x_h) in occupation o at the beginning of next period's separation stage is given by

$$\begin{aligned} e^s_{t+1}(z, x_h, o)dz &= (1-\mu)\Bigg[\chi^e(x_h|x_h)\int_{\underline{z}}^{\overline{z}} e^p_t(\tilde{z}, x_h, o)dF(z|\tilde{z})d\tilde{z} \\ &+ (\mathbf{1}_{h>1})\chi^e(x_h|x_{h-1})\int_{\underline{z}}^{\overline{z}} e^p_t(\tilde{z}, x_{h+1}, o)dF(z|\tilde{z})d\tilde{z}\Bigg]. \end{aligned} \tag{16}$$

Conditional on not retiring from the labor market, the terms inside the squared bracket show the probability that employed workers in labor markets $(\tilde{z}, x_h, o)$ and $(\tilde{z}, x_{h-1}, o)$ in the previous period's production stage will be in labor market (z, x_h, o) immediately after the z and x_h shocks occur. In this case, the indicator function $\mathbf{1}_{h>1}$ takes the value of one when the labor market (z, x_h) is associated with a value of $x_h > x_1$ and zero otherwise.

Since we count those employed workers who separated from their employers in the production stage and employed workers do not participate in the reallocation or the search and matching stages, it follows that $e^s_{t+1}(z, x_h, o)dz = e^r_{t+1}(z, x_h, o)dz = e^m_{t+1}(z, x_h, o)dz$. This implies that the measure of employed workers characterised by (z, x_h) in occupation o during the production stage is given by

$$e^p_{t+1}(z, x_h, o)dz = (1 - d(z, x_h, o, A, \mathcal{P}_O))e^s_{t+1}(z, x_h, o)dz + \lambda(\theta(\omega))u^m_{t+1}(z, x_h, o)dz, \qquad (17)$$

where the last term describes those unemployed workers in labor market (z, x_h) who found a job in their same occupation o.

B.3 Proofs

Definition A Block Recursive Equilibrium (BRE) is a set of value functions $W^U(\omega)$, $W^E(\omega)$, $J(\omega)$, workers' policy functions $d(\omega)$, $\rho(\omega)$, $\mathcal{S}(\omega)$, firms' policy function $\sigma(\omega)$, tightness function $\theta(\omega)$, wages $w(\omega)$, laws of motion of A, $\mathcal{P}_O$, z and x for all occupations, and laws of motion for the distribution of unemployed and employed workers over all occupations, such that: *(i)* the value functions and decision rules follow from the firm's and worker's problems described in equations (1)-(5) in the main text; *(ii)* labor market tightness $\theta(\omega)$ is consistent with free entry on each labor market, with zero expected profits determining $\theta(\omega)$ on labor markets at which positive ex-post profits exist; $\theta(\omega) = 0$ otherwise; *(iii)* wages solve equation (6) in the main text; *(iv)* the worker flow equations map initial distributions of unemployed and employed workers (respectively) over labor markets and occupations into next period's distribution of unemployed and employed workers over labor markets and occupations, according to the above policy functions and exogenous separations.

Proposition 2 Given $F(z'|z) < F(z'|\tilde{z})$ for all z, z' when $z > \tilde{z}$: (i) a BRE exists and it is the unique equilibrium, and (ii) the BRE is constrained efficient.

B.3.1 Proof of Proposition 2

We divide the proof into two parts. In the first part we show existence of equilibrium by deriving the operator T, showing it is a contraction and then verifying that the candidate equilibrium functions from the fixed point of T satisfy all equilibrium conditions. The second part shows efficiency of equilibrium.

Existence

Step 1: Let $M(\omega) \equiv W^E(\omega) + J(\omega)$ denote the value of the match. We want to show that the value functions $M(\omega)$, $W^U(\omega)$ and $R(\omega)$ exist. This leads to a three dimensional fixed point problem. It is then useful to define the operator T that maps the value function $\Gamma(\omega, n)$ for $n = 0, 1, 2$ into the same functional space, such that $\Gamma(\omega, 0) = M(\omega)$, $\Gamma(\omega, 1) = W^U(\omega)$, $\Gamma(\omega, 2) = R(\omega)$ and

$$T(\Gamma(\omega, 0)) = y(A, p_o, z, x) + \beta \mathbb{E}_{\omega'}\Big[\max_{d^T}\{(1 - d^T)M(\omega') + d^T W^U(\omega')\}\Big],$$

$$T(\Gamma(\omega, 1)) = b + \beta \mathbb{E}_{\omega'}\Big[\max_{\rho^T}\{\rho^T R(\omega') + (1 - \rho^T)(D^T(\omega') + W^U(\omega'))\}\Big],$$

$$T(\Gamma(\omega, 2)) = \max_{\mathcal{S}(\omega)}\Big(\sum_{\tilde{o} \in \vec{O}^-} \alpha(s^T_{\tilde{o}}) \int_{\underline{z}}^{\overline{z}} W^U(\tilde{z}, x_1, \tilde{o}, A, \mathcal{P}_O)dF(\tilde{z}) + (1 - \sum_{\tilde{o} \in \vec{O}^-} \alpha(s^T_{\tilde{o}}))[b + \beta \mathbb{E}_{\omega'} R(\omega')] - c\Big),$$

where the latter maximization is subject to $s_o \in [0,1]$ and $\sum_{o\in\vec{O}^-} s_o = 1$, and

$$D^T(\omega') \equiv \lambda(\theta(\omega'))(1-\eta)\Big(M(\omega') - W^U(\omega')\Big), \text{ with } \theta(\omega') = \left(\frac{\eta(M(\omega') - W^U(\omega'))}{k}\right)^{\frac{1}{1-\eta}}. \quad (18)$$

Lemma A.1: T is (i) a well-defined operator mapping functions from the closed space of bounded continuous functions $\mathcal{E}$ into $\mathcal{E}$, and (ii) a contraction.

First we show that the operator T maps bounded continuous functions into bounded continuous functions. Let $W^U(\omega)$, $R(\omega)$ and $M(\omega)$ be bounded continuous functions. It then follows that $\lambda(\theta(\omega))$ and $D(\omega)$ are continuous functions. It also follows that $\max\{M(\omega), W^U(\omega)\}$ and $\max\{R(\omega), D(\omega)+W^U(\omega)\}$ are continuous. Further, since the constraint set for $\mathcal{S}$ is compact-valued and does not depend on ω, functions $\alpha(.)$ are continuous, and the integral of continuous function W^U is continuous, the theorem of the maximum then implies that the expression

$$\max_{\mathcal{S}(\omega)}\Big(\sum_{\tilde{o}\in\vec{O}^-}\alpha(s^T_{\tilde{o}})\int_{\underline{z}}^{\overline{z}} W^U(\tilde{z}, x_1, \tilde{o}, A, \mathcal{P}_O)dF(\tilde{z}) + (1-\sum_{\tilde{o}\in\vec{O}^-}\alpha(s^T_{\tilde{o}}))[b+\beta\mathbb{E}_{\omega'}R(\omega')] - c\Big) \quad (19)$$

is continuous. Therefore T maps continuous functions into continuous functions. Moreover, since the domain of ω is bounded, and $\alpha(.)$ and $\lambda(.)$ are bounded on bounded domains, T maps the space of bounded continuous functions into itself.

Second we show that T defines a contraction. Consider two functions $\Gamma, \Gamma' \in \mathcal{E}$, such that $\|\Gamma - \Gamma'\|_{\text{sup}} < \varepsilon$. It then follows that $\|W^U - W^{U\prime}\|_{\text{sup}} < \varepsilon$, $\|R - R'\|_{sup} < \varepsilon$ and $\|M - M'\|_{\text{sup}} < \varepsilon$, where W^U, R and M are part of Γ as defined above. We first establish that

$$\|D + W^U - D' - W^{U\prime}\|_{sup} < \varepsilon, \quad (20)$$

when function-tuples D and D' are derived from (M, W^U), and $(M', W^{U\prime})$ respectively. With this aim consider, without loss of generality, the case in which $M(\omega) - W^U(\omega) > M'(\omega) - W^{U\prime}(\omega)$ at a given ω. Instead of $D(\omega)$, write $D(M(\omega) - W(\omega))$ to make explicit the dependence of D on M and W. To reduce notation we suppress the dependence on ω for this part of the proof and further condense $W^U = W$. From $M - W > M' - W'$, it follows that $\varepsilon > W' - W \geq M' - M > -\varepsilon$. Construct $M'' = W' + (M - W) > M'$ and $W'' = M' - (M - W) < W'$. Equation (18) implies $D(M - W) + W$ is increasing in M and in W. To verify this, hold M constant and note that

$$\frac{d(D(M-W)+W)}{dW} = -\lambda(\theta(.)) + 1 \geq 0 \quad (21)$$

by virtue of $d(D(M-W))/d(M-W) = \lambda$ and that $\lambda \in [0,1]$, where the inequality in (21) is strict when $\lambda \in [0,1)$ and weak when $\lambda = 1$. Then, it follows that

$$\begin{aligned}-\varepsilon < D(M' - W'') + W'' - D(M - W) - W &\leq D(M' - W') + W' - D(M - W) - W\\ &\leq D(M'' - W') + W' - D(M - W) - W < \varepsilon\end{aligned}$$

where $D(M' - W'') = D(M - W) = D(M'' - W')$ by construction. Note that the outer inequalities follow because $M - M' > -\varepsilon$ and $W' - W < \varepsilon$. Given that D, M and W are bounded continuous functions on a compact domain and the above holds for every ω, it then must be that $\|D + W^U - D' - W^{U'}\|_{\text{sup}} < \varepsilon$. Since $\|\max\{a,b\} - \max\{a', b'\}\| < \max\{\|a - a'\|, \|b - b'\|\}$, as long as the terms over which to maximize do not change by more than ε in absolute value, the maximized value does not change by more ε, it then follows that $\|T(\Gamma(\omega, n)) - T(\Gamma'(\omega, n)\| < \beta\varepsilon$ for all ω; $n = 0, 1$.

To show that $\|T(\Gamma(\omega, 3)) - T(\Gamma'(\omega, 3)\| < \beta\varepsilon$ for all ω, with a slight abuse of notation, define $TR(\omega, \mathcal{S})$ as

$$TR(\omega, \mathcal{S}) = \sum_{\tilde{o} \in \vec{O}^-} \alpha(s_{\tilde{o}}(\omega)) \int_{\underline{z}}^{\overline{z}} \Big[b + \beta \mathbb{E}_{\omega'} \Big[\max_{\rho^T} \{ \rho^T R(\omega') + (1 - \rho^T)(D^T(\omega') + W^U(\omega')) \} \Big] \Big] dF(\tilde{z}) \tag{22}$$

$$+ (1 - \sum_{\tilde{o} \in \vec{O}^-} \alpha(s_{\tilde{o}}(\omega)))[b + \beta \mathbb{E}_{\omega'} R(\omega')] - c.$$

Let $\mathcal{S}_*(\omega) = \{s_1^*, \ldots, s_{o-1}^*, s_{o+1}^*, \ldots, s_O^*\}$ be the maximizer of TR at ω and $\mathcal{S}_*'(\omega)$ of TR'. Without loss of generality assume that $TR(\omega, \mathcal{S}) > TR'(\omega, \mathcal{S}')$. Since our previous arguments imply that

$$\| \Big[\max_{\rho} \{ \rho R + (1 - \rho)(D + W^U) \} \Big] - \Big[\max_{\rho} \{ \rho R' + (1 - \rho)(D' + W') \} \Big] \|_{sup} < \beta\varepsilon$$

it then follows that $\|TR(\omega, \mathcal{S}_*) - TR'(\omega, \mathcal{S}_*)\|_{sup} < \beta\varepsilon$. Further, since $TR(\omega, \mathcal{S}_*) > TR'(\omega, \mathcal{S}_*') > TR'(\omega, \mathcal{S}_*)$ we have that $0 < TR(\omega, \mathcal{S}_*) - TR'(\omega, \mathcal{S}_*) < \beta\varepsilon$. Using the latter inequality to obtain $0 < TR(\omega, \mathcal{S}_*) - TR'(\omega, \mathcal{S}_*') + TR'(\omega, \mathcal{S}_*') - TR'(\omega, \mathcal{S}_*) < \beta\varepsilon$ and noting that $TR'(\omega, \mathcal{S}_*') - TR'(\omega, \mathcal{S}_*) > 0$, it follows that $0 < TR(\omega, \mathcal{S}_*) - TR'(\omega, \mathcal{S}_*') < \beta\varepsilon$. This last step is valid for all ω, then it implies that $\|T(\Gamma(\omega, 3)) - T(\Gamma'(\omega, 3)\|_{sup} < \beta\varepsilon$ and hence the operator T is a contraction and has a unique fixed point.

Step 2 (Linking the Mapping T and BRE Objects): From the fixed point functions $M(\omega)$, $W^U(\omega)$ and $R(\omega)$ define the function $J(\omega) = \max\{(1 - \zeta)[M(\omega) - W^U(\omega)], 0\}$, and the functions $\theta(\omega)$ and $V(\omega)$ from $0 = V(\omega) = -k + q(\theta(\omega))J(\omega)$. Also define $W^E(\omega) = M(\omega) - J(\omega)$ if $M(\omega) > W^U(\omega)$, and $W^E(\omega) = M(\omega)$ if $M(\omega) \leq W^U(\omega)$. Finally, define $d(\omega) = d^T(\omega)$, $\sigma(\omega) = \sigma^T(\omega)$, $\rho(\omega) = \rho^T(\omega)$, $\mathcal{S}(\omega) = \mathcal{S}^T(\omega)$ and a $w(\omega)$ derived using the Nash bargaining equation in the main text given all other functions.

Given $1 - \zeta = \eta$ and provided that the job separation decisions between workers and firms coincide, which they are as a match is broken up if and only if it is bilaterally efficient to do so according to $M(\omega)$ and $W^U(\omega)$, then equations (5) and (6) in the main text (describing $J(\omega)$ and surplus sharing) are satisfied. Further, equation (3) in the main text (describing $W^E(\omega)$) is satisfied by construction, $\theta(\omega)$ satisfies the free-entry condition and $w(\omega)$ satisfies equation (6) in the main text. Hence, the constructed value functions and decision rules satisfy all conditions of the equilibrium and the implied evolution of the distribution of employed and unemployed workers also satisfies the equilibrium conditions.

Uniqueness follows from the same procedure in the opposite direction and using a contradiction argument. Suppose the BRE is not unique. Then a second set of functions exists that satisfy all the equilibrium conditions. Construct $\hat{M}$, $\hat{W}^U$ and $\hat{R}$ from these conditions. Since in any equilibrium the job separation decisions have to be bilaterally efficient and the occupational mobility decisions (ρ and $\mathcal{S}$) are captured in T, then $\hat{M}$, $\hat{W}^U$ and $\hat{R}$ must be a fixed point of T, contradicting the uniqueness of the fixed point established by Banach's Fixed Point Theorem. Hence, there is a unique BRE.

Efficiency

The social planner, currently in the production stage, solves the problem of maximizing total discounted output by choosing job separations decisions $d(.)$, occupational mobility decisions $\rho(.)$ and $\mathcal{S}(.)$, as well as vacancy creation decisions $v(.)$ for each pair (z, x_h) across all occupations $o \in O$ in any period t. The first key aspect of the planner's choices is that they could potentially depend on the entire state space $\Omega^j = \{z, x, o, A, \mathcal{P}_O, \mathcal{G}^j\}$ for each of the four within period stages $j = s, r, m, p$ (separation, reallocation, search and matching, production) and workers' employment status, such that its maximization problem is given by

$$\max_{\{d(\Omega^s),\rho(\Omega^r),\mathcal{S}(\Omega^r),v(\Omega^m)\}} \mathbb{E}\sum_{t=0}^{\infty}\Bigg(\sum_{o=1}^{O}\sum_{h=1}^{H}\int_{\underline{z}}^{\overline{z}}\beta^t\left[u^p_t(z,x_h,o)b+e^p_t(z,x_h,o)y(A_t,p_{o,t},z,x_h)\right]dz \tag{23}$$
$$-\sum_{o'=1}^{O}\sum_{h'=1}^{H}\int_{\underline{z}}^{\overline{z}}\beta^{t+1}\Big[c\rho(z',x'_h,o',A_{t+1},\mathcal{P}_{O,t+1},\mathcal{G}^r_{t+1})u^r_{t+1}(z',x'_h,o')$$
$$+kv_{t+1}(z',x'_h,o',A_{t+1},\mathcal{P}_{O,t+1},\mathcal{G}^m_{t+1})\Big]dz'\Bigg),$$

subject to the initial conditions $(A_0, \mathcal{P}_{O,0}, \mathcal{G}^p_0)$, the laws of motion for unemployed and employed workers described in Section 2 of this appendix (with corresponding state space Ω^j for the decision rules), and the choice variables $\rho(.)$ and $d(.)$ being continuous variables in $[0, 1]$, as the planner can decide on the proportion of workers in labor market (z, x_h) to separate from their jobs or to change occupations.

Note that implicitly the social planner is constrained in the search technology across occupations: it faces the same restrictions as an individual worker (in occupation $o \in O$), on the proportion of time that can be devoted to obtain a z-productivity from occupation $\tilde{o} \neq o$. Namely, $s_{\tilde{o}}(.) \in [0, 1]$, $\sum_{\tilde{o}\in\vec{O}^-_o} s_{\tilde{o}}(.) = 1$ and $\sum_{\tilde{o}\in\vec{O}^-_o} \alpha(s_{\tilde{o}}(.), o) \leq 1$, where $\vec{O}^-_o$ denotes the set of remaining occupations relative to o. The latter notation highlights that, as in the decentralised problem, once the occupational mobility decision has been taken the new z-productivity cannot be obtain from the departing occupation.

Rewriting the planners' problem in recursive form as the fixed point of the mapping T^{SP} and letting next period's values be denoted by a prime yields

$$T^{SP}W^{SP}(\Omega^p) = \max_{\substack{\{d(\Omega^{s'}),\rho(\Omega^{r'}),\\ \mathcal{S}(\Omega^{r'}),v(\Omega^{m'})\}}} \sum_{o=1}^{O}\sum_{h=1}^{H}\int_{\underline{z}}^{\overline{z}}(u^p(z,x_h,o)b+e^p(z,x_h,o)y(A,p_o,z,x_h))dz \tag{24}$$
$$+\beta\mathbb{E}_{\Omega^{s\prime}|\Omega^p}\left[-\left(c\sum_{o=1}^{O}\sum_{h=1}^{H}\int_{\underline{z}}^{\overline{z}}\rho(\Omega^{r'})u^{r\prime}(z',x'_h,o)dz'+k\sum_{o=1}^{O}\sum_{h=1}^{H}\int_{\underline{z}}^{\overline{z}}v(\Omega^{m'})dz'\right)+W^{SP}(\Omega^{p\prime})\right],$$

subject to the same restrictions described above. Our aim is to show that this mapping is a contraction that maps functions $W^{SP}(.)$ from the space of functions linear with respect to the distribution $\mathcal{G}^p$ into itself, such that

$$W^{SP}(\Omega^p) = \sum_{o=1}^{O}\sum_{h=1}^{H}\int_{\underline{z}}^{\overline{z}}\Big(u^p(z,x_h,o)W^{u,SP}(\omega)+e^p(z,x_h,o)M^{SP}(\omega)\Big)\,dz, \tag{25}$$

for some functions $W^{u,SP}(\omega)$ and $M^{SP}(\omega)$, where $\omega = (z, x_h, o, A, \mathcal{P}_o)$, and $u^p(.)$ and $e^p(.)$ are implied by $\mathcal{G}^p$.

In Section 2 of this appendix, we decomposed the next period's measure of unemployed workers at the production stage, $u^{p'}$, into three additive terms (see equation (15)). The first term refers to those unemployed workers that were unsuccessful in matching, either due to not finding a posted vacancy in their submarket or because the planner chose not to post vacancies in their submarket. The second term refers to those workers who separated from employment and hence were restricted from the search and matching stage within the period. The third term refers to those who changed occupations and came from other markets into submarket (z, x_h, o) and were restricted from the search and matching stage within the period. Likewise, we decomposed the next period's measure of employed workers at the production stage into two additive terms (see equation (17)). The first one refers to the survivors in employment from the previous separation stage, while the second terms refers to new hires. Given a continuation value function $W^{SP}(\Omega^{p'})$ that is linear in $\mathcal{G}^p$, as in (25), we now show that we can also decompose the expression in (24) into different additive components.

Consider first those unemployed workers who are eligible to participate in the search and matching stage. Note that the matching technology implies $v(\Omega^m) = \theta(\Omega^m)(1-\rho(\Omega^r))u^m(z, x_h, o)$. Therefore we can consider $\theta(\Omega^{m'})$ as the planner's choice in equation (24) instead of $v(\Omega^{m'})$. Next we isolate the terms of $W^{SP}(\Omega^{p'})$ that involve workers who went through the search and matching stage: the first term of the unemployed worker flow equation (15) and the second term of the employed worker flow equation (17). We then combine these terms with the cost of vacancy posting in (24). Noting that the dependence of $u^m(\Omega^{m'})$ captures a potential dependence of the planner's occupational mobility decisions made in the previous reallocation stage, we can express the terms in the mapping (24) that involve $\theta(\Omega^{m'})$ as

$$\sum_{o=1}^{O}\sum_{h=1}^{H}\int_{\underline{z}}^{\overline{z}} u^{m\prime}(\Omega^{m'})\Big\{-k\theta(.) + \lambda(\theta(.))M^{SP}(\omega') + (1-\lambda(\theta(.))W^{u,SP}(\omega')\Big\}dz'. \tag{26}$$

Note that by backwards induction, the planner's optimal decisions for $\theta(\Omega^{m'})$ maximize (26). Given that the terms within the curly brackets only depend on ω' when the continuation value is linear as in (25), maximizing with respect to θ implies an optimal $\theta(\omega')$. Using this result we let $W^{m,SP}(\omega')$ denote the sum of the maximized terms inside the curly brackets in (26) .

Next consider the search-direction choice across occupations for those workers who the planner has decided to move across occupations. The following expression summarizes the terms of (24) that involve this choice,

$$\sum_{o'}\sum_{h'=1}^{H}\int_{\underline{z}}^{\overline{z}}\Big[u^r_{t+1}(\Omega^{r'})\rho(\Omega^{r'})\Big\{\sum_{\hat{o}\neq o'}\alpha(s_{\hat{o}}(.), o')\int_{\underline{z}}^{\overline{z}}\Big[W^{u,SP}(\hat{z}, x_1, \hat{o}, A', \mathcal{P}'_O)\Big]dF(\hat{z})$$
$$+(1-\sum_{\hat{o}\neq o'}\alpha(s_{\hat{o}}(.), o'))[b + \beta\mathbb{E}_{\omega''}R^{SP}(\omega'')]\Big\}\Big]dz', \tag{27}$$

where we have applied the properties of the z-productivity process to change the order of integration relative to the third term of flow equation (15) and have integrated over all of next period's production stage states. Since we allow previous decisions to depend on the entire state space, $u^r_{t+1}(.)$ and $\rho(.)$ are written as a function of $\Omega^{r'}$ as they potentially depend on the distribution $\mathcal{G}^r$. Note, however, that the term inside the curly brackets in (27) can be maximized separately for each $(z', {x_h}', o')$ and can therefore be summarized by $R^{gross,SP}(\omega') = R^{SP}(\omega') - c$, while the search intensity decision $\mathcal{S}(.)$ has solution vector $\{s(\omega', \tilde{o})\}$.

Noting that $u^s(.) = u^r(.)$ and $u^m(.) = (1-\rho(.))u^r(.)$ (see Section 2 of this appendix), the above results imply we can express the term on the second line of equation (24) as

$$\beta \mathbb{E}_{\Omega^{s'}|\Omega^p}\Bigg[\Bigg(\sum_{o=1}^{O}\sum_{h=1}^{H}\int_{\underline{z}}^{\overline{z}}\Big[u^{s\prime}(z', x_h', o)\Big\{\rho(.)R^{SP}(\omega') + (1-\rho(.))W^{m,SP}(\omega')\Big\} \qquad (28)$$

$$+ e^{s\prime}(z', x_h', o)\Big\{(1-d(.))M^{SP}(\omega') + d(.)W^{u,SP}(\omega')\Big\}\Big]dz\Bigg)\Bigg],$$

where it follows that the maximizing decisions $\rho(.)$ and $d(.)$ are also functions of ω', as these are the only other dependencies within the curly brackets.

Finally, given the properties of the shock processes, note that $u^s_{t+1}(.)$ and $e^s_{t+1}(.)$ are linear functions of $u^p(.)$ and $e^p(.)$, as demonstrated by equations (14) and (16) in Section 2 of this appendix. This implies that we can express $T^{SP}W^{SP}(\Omega^p)$ as

$$T^{SP}W^{SP}(\Omega^p) = \sum_{o=1}^{O}\sum_{h=1}^{H}\int_{\underline{z}}^{\overline{z}}\left[TW^{U,SP}_{max}(z, x_h, o, A, \mathcal{P}_O)u(z, x_h, o) + TM^{SP}_{max}(z, x_h, o, A, \mathcal{P}_O)e(z, x_h, o)\right]dz, \qquad (29)$$

where $TW^{U,SP}_{max}$ is given by

$$TW^{U,SP}_{max}(\omega) = \max_{\rho(\omega'),\theta(\omega')}\Bigg\{b + \beta\mathbb{E}_{\omega'|\omega}\Bigg[\rho(\omega')\Bigg(\int_{\underline{z}}^{\overline{z}}\max_{\mathcal{S}(\omega')}\Big[\sum_{\tilde{o}\neq o}\alpha(s_{\tilde{o}}(\omega', o))W^{U,SP}_{max}(\tilde{z}, x_1, \tilde{o}, A', \mathcal{P}'_O)dF(\tilde{z})$$

$$+ (1 - \sum_{\tilde{o}\neq o}\alpha(s_{\tilde{o}}(\omega', o)))[b + \beta\mathbb{E}_{\omega''|\omega'}R^{U,SP}_{max}(\omega'')\Big] - c\Bigg) \qquad (30)$$

$$+ (1-\rho(\omega'))\Big[\lambda(\theta(\omega'))(M^{SP}_{max}(\omega') - W^{U,SP}_{max}(\omega')) - \theta(\omega')k + W^{U,SP}_{max}(\omega')\Big]\Bigg]\Bigg\},$$

and TM^{SP}_{max} is given by

$$TM^{SP}_{max}(p, x, z) = \max_{d(\omega')}\Bigg\{y(z, x_h, A, p_o) + \beta\mathbb{E}_{\omega'|\omega}\Bigg[\Big(d(\omega')W^{U,SP}_{max}(\omega') + (1-d(\omega'))M^{SP}_{max}(\omega')\Big)\Bigg]\Bigg\}. \qquad (31)$$

Hence we have established that the mapping T^{SP} described in equation (24) maps functions $W^{SP}(.)$ from the space of functions linear with respect to the distribution $\mathcal{G}^p$ (of the form (25)) into itself.

It is now straightforward to show that if $TW^{U,SP}_{max}$, TR^{SP}_{max} and TM^{SP}_{max} are contraction mappings, then (29) (and thereby (24)) is also a contraction mapping. Given the regularity properties assumed on the shock processes and following the proof of the decentralised case in Section 3.1.1 of this appendix, we can show that a fixed point exists for (30) and (31). Using equation (29), we then can construct the fixed point of the expression in (24). It then follows from (30) and (31) that if the Hosios' condition holds, allocations of the fixed point of T are allocations of the fixed point of T^{SP}, and hence the equilibrium allocations in the decentralised setting are also the efficient allocations.

BRE gives the *unique* equilibrium allocation

We now show that in any equilibrium, decisions and value functions only depend on $\omega = (z, x_h, o, A, \mathcal{P}_o)$. We proceed using a contradiction argument. Suppose there is an alternative equilibrium in which values and decisions do not depend only on ω, but also on an additional factor like the entire distribution of workers over employment status and z-productivities $\mathcal{G}$, or its entire history of observables, H_t. Consider the associated value functions in this alternative equilibrium, where the relevant state vector of the alternative equilibrium is given by (ω, H_t). Our aim is to show that such an equilibrium cannot exist.

First suppose that in the alternative equilibrium all value functions are the same as in the BRE, but decisions differ at the same ω. This violates the property that, in our setting, all maximizers in the BRE value functions are unique, leading to a contradiction. Now suppose that in the alternative equilibrium at least one value function differs from the corresponding BRE value function at the same ω. It is straightforward to show that the expected values of unemployment must differ in both equilibria. Let $W^U(\omega, H_t)$ denote the value function for unemployed workers in the alternative equilibrium, and let $W^U(\omega)$ denote the corresponding value function in the BRE for the same ω. Since in the proof of efficiency of a BRE we did not rely on the uniqueness of the BRE in the broader set of all equilibria, we can use the proved results of Section 3.1.2 of this appendix here.

In particular, recall that the social planner's problem is entirely linear in the distribution of workers across states and hence $W^U(\omega)$ is the best the unemployed worker with ω can do (without transfers), including in the market equilibrium. Likewise, $M(\omega)$ is the highest value of the joint value of a match, including in the market equilibrium. Since value functions are bounded from above and from below and are continuous in their state variables, there exists a supremum of the difference between $W^U(\omega)$ and the candidate market equilibrium's $W^U(\omega, H_t)$, $\sup(W^U(\omega) - W^U(\omega, H_t)) = \epsilon_u > 0$. Similarly, there also exists a supremum for the difference between $M(\omega)$ and $M(\omega, H_t)$, $\sup(M(\omega) - M(\omega, H_t)) = \epsilon_m > 0$. In what follows, we will show that a difference in the value functions for unemployed workers (or the value functions for the joint value of the match) arbitrarily close to $\epsilon_u > 0$ (or $\epsilon_m > 0$) cannot occur. Otherwise, this will require that the difference in *tomorrow's* values to be larger than ϵ_u (or ϵ_m). In turn, this implies that an alternative equilibrium cannot exist.

From the above definition of supremum, it follows that

$$\max\{M(\omega), W^U(\omega)\} - \max\{M(\omega, H_t), W^U(\omega, H_t)\} < \max\{\epsilon_u, \epsilon_m\}.$$

Since

$$M(\omega) = y(z, x_h, A, p_o) + \beta\mathbb{E}[\max\{M(\omega'), W^U(\omega')\}],$$

and likewise for $M(\omega, H_t)$ it follows that at any (ω, H_t),

$$M(\omega) - M(\omega, H_t)) < \beta \max\{\epsilon_u, \epsilon_m\}. \tag{32}$$

Consider first the case in which $\epsilon_m \geq \epsilon_u$. For (ω, H_t) achieving a difference $M(\omega) - M(\omega, H_t) > \beta\epsilon_m$ is not possible since this will lead to a contradiction in equation (32) when $\epsilon_m > 0$.

Next consider the case in which $\epsilon_m < \epsilon_u$. Simplifying notation by dropping ω and using the prime instead of (ω, H_t), we first establish an intermediate step. At *any* (ω, H_t) it holds that

$$\lambda(\theta)(1-\eta)M + (1-\lambda(\theta)(1-\eta))W < \lambda(\theta')(1-\eta)M' + (1-\lambda(\theta')(1-\eta))W' + \epsilon_u. \quad (33)$$

There are two cases to be analysed to show the above relationship.

Case 1: Suppose that $(M'-W') \geq M-W$, then $\lambda(\theta') \geq \lambda(\theta)$. Define $K = (1-\eta)(\lambda(\theta') - \lambda(\theta))(M'-W') \geq 0$. Combining the latter with $\lambda(\theta)(1-\eta)(M-M')+(1-\lambda(\theta)(1-\eta))(W-W') \leq \epsilon_u$, it must be true that

$$\lambda(\theta)(1-\eta)(M-M') + (1-\lambda(\theta)(1-\eta))(W-W') - K \leq \epsilon_u, \quad (34)$$

from which (33) follows.

Case 2: Suppose that $(M-W) > (M'-W')$, then $\lambda(\theta) > \lambda(\theta')$. From the derivative of $\frac{d}{d(M-W)}(\lambda(\theta(M-W))(1-\eta)(M-W)) = \lambda(\theta(M-W))$, we can establish that, if $(M-W) > (M'-W')$,

$$\lambda(\theta)((M-W)-(M'-W')) > \lambda(\theta)(1-\eta)(M-W)-\lambda(\theta')(1-\eta)(M'-W') > \lambda(\theta')((M-W)-(M'-W')).$$

Since in any equilibrium (not only in the BRE) θ' depends only on $(M'-W')$ and constant parameters, we can use this relationship to establish that

$$W + \lambda(\theta)(M-W) - (W' + \lambda(\theta)(M'-W')) < \epsilon_u,$$

from which (33) follows.

The final step is to consider an (ω, H_t) such that $\max\{\epsilon_m, \beta^{-1}\epsilon_u\} < W^U(\omega) - W^U(\omega, H_t) < \epsilon_u$, where such a (ω, H_t) exists by the definition of supremum. With this in hand it is straightforward to check that the difference in *tomorrow*'s value (under the expectation sign), between $W^U(\omega)$ and $W^U(\omega, H_t)$ will not exceed ϵ_u, since term-by-term, the difference is bounded by ϵ_u. This also implies that today's difference, $W^U(\omega) - W^U(\omega, H_t)$, cannot be more than $\beta\epsilon_u > 0$, which contradicts our premise. This establishes that a difference in the value functions for unemployed workers (or the value functions for the joint value of the match) arbitrarily close to $\epsilon_u > 0$ (or $\epsilon_m > 0$) cannot occur. Hence the BRE is the unique equilibrium.

This completes the proof of Proposition 2.

B.3.2 Proof of existence of a reallocation and separation cutoff

Reservation property of occupational mobility decisions, z^r Here we show that $M(\omega)$ and $W^U(\omega)$ as derived in the proof of Proposition 2 are increasing in z. If $M(\omega)$ and $W^U(\omega)$ are continuous and bounded functions increasing in z, T maps them into increasing (bounded and continuous) functions. For employed workers $(T(\Gamma(\omega, 0))$, this follows since both $\max\{M(\omega'), W^U(\omega')\}$ and $y(.)$ are increasing in z, while the stochastic dominance of the z-productivity transition law implies higher expected z's tomorrow. For unemployed workers $(T(\Gamma(\omega, 1))$, note that the value of changing occupations $R(\omega)$ does not depend on the current z of the worker, while equation (21) implies that $D(M(\omega) - W^U(\omega)) + W^U(\omega)$ is increasing in z. Again, given stochastic dominance of the tomorrow's z when today's z is higher, $T(\Gamma(\omega, 1))$ is also increasing in z. The reservation property follows immediately, since $R(\omega)$ is constant in z and $D(M(\omega) - W^U(\omega)) + W^U(\omega)$ is increasing in z.

Reservation property of job separation decisions, z^s We now show that $M(\omega) - W^U(\omega)$ is increasing in z when $\delta + \lambda(\theta(\omega)) < 1$ for the case of no human capital accumulation and occupational-wide shocks. In the calibration we show that this property holds also for the case of human capital accumulation and occupational-wide shocks.

Consider the same operator T defined in the proof of Proposition 2, but now the relevant state space is given by (A, z). Note that the value functions describing the worker's and the firm's problem, do not change, except for the fact that we are using a smaller state space. It is straightforward to verify that the derived properties of T in Lemma A.1 also apply in this case. We now want to show that this operator maps the subspace of functions Γ into itself with $M(A, z)$ increasing weakly faster in z than $W^U(A, z)$. To show this, take $M(A, z)$ and $W^U(A, z)$ such that $M(A, z) - W^U(A, z)$ is weakly increasing in z and let z^s denote a reservation productivity such that for $z < z^s$ a firm-worker match decide to terminate the match. Using $\lambda(\theta)(M - W^U) - \theta k = \lambda(\theta)(M - W^U) - \lambda'(\theta)(M - W^U)\theta = \lambda(\theta)(1 - \eta)(M - W^U)$, we construct the following difference

$$T\Gamma(A, z, 0) - T\Gamma(A, z, 1) = \tag{35}$$
$$y(A, z) - b + \beta\mathbb{E}_{A,z}\Big[(1 - \delta)\max\{M(A', z') - W^U(A', z', 0\} -$$
$$\max\Big\{\int W^U(A', \tilde{z})dF(\tilde{z}) - c - W^U(A', z'), \lambda(\theta)(1 - \eta)\big(M(A', z') - W^U(A', z')\big)\Big\}\Big].$$

The first part of the proof shows the conditions under which $T\Gamma(A, z, 0) - T\Gamma(A, z, 1)$ is weakly increasing in z. Because the elements of the our relevant domain are restricted to have $W^U(A, z)$ increasing in z, and $M(A, z) - W^U(A, z)$ increasing in z, we can start to study the value of the term under the expectation sign, by cutting a number of different cases to consider depending on where z' is relative to the implied reservation cutoffs.

– *Case 1.* Consider the range of tomorrow's $z' \in [\underline{z}(A'), z^r(A'))$, where $z^r(A') < z^s(A')$. In this case, the term under the expectation sign in the above equation reduces to $-\int W^U(A', \tilde{z})dF(\tilde{z}) + c + W^U(A', z')$, which is increasing in z'.

– *Case 2.* Now suppose tomorrow's $z' \in [z^r(A'), z^s(A'))$. In this case, the term under the expectation sign becomes zero (as $M(A', z') - W^U(A', z') = 0$), and is therefore constant in z'.

– *Case 3.* Next suppose that $z' \in [z^s(A'), z^r(A'))$. In this case, the entire term under the expectation sign reduces to

$$(1 - \delta)(M(A', z') - W^U(A', z')) - \int W^U(A', \tilde{z})dF(\tilde{z}) + c + W^U(A', z'),$$

and, once again, is weakly increasing in z', because by supposition $M(A', z') - W^U(A', z')$ is weakly increasing in z', and so is $W^U(A', z')$ by Lemma A.1.

– *Case 4.* Finally consider the range of $z' \geq \max\{z^r(A'), z^s(A')\}$, such that in this range employed workers do not quit nor reallocate. In this case the term under the expectation sign equals

$$(1 - \delta)[M(A', z') - W^U(A', z')] - \lambda(\theta(A', z'))(1 - \eta)[M(A', z') - W^U(A', z')]. \tag{36}$$

It is easy to show using the free entry condition that $\frac{d(\lambda(\theta^*(A',z'))(1-\eta)[M(A',z')-W^U(A',z')])}{d(M-W)} = \lambda(\theta(A', z'))$, and hence that the derivative of (36) with respect to z' is positive whenever $1 - \delta - \lambda(\theta) \geq 0$.

Given $F(z'|z) < F(z'|\tilde{z})$ for all z, z' when $z > \tilde{z}$, the independence of z of A, and that the term under the expectation sign are increasing in z', given any A', it follows that the integral in (35) is increasing in today's z. Together with $y(A, z)$ increasing in z, it must be that $T\Gamma(A, z, 1) - T\Gamma(A, z, 0)$ is also increasing in z.

To establish that the fixed point also has increasing differences in z between the first and second coordinate, we have to show that the space of this functions is closed in the space of bounded and continuous functions. In particular, consider the set of functions $\mathbb{F} \stackrel{def}{=} \{f \in \mathcal{C} | f : X \times Y \to \mathbb{R}^2, |f(x, y, 1) - f(x, y, 2)| \text{ increasing in } y\}$, where $f(.,.,1), f(.,.,2)$ denote the first and second coordinate, respectively, and $\mathcal{C}$ the metric space of bounded and continuous functions endowed with the sup-norm.

The next step in the proof is to show that fixed point of $T\Gamma(A, z, 0) - T\Gamma(A, z, 1)$ is also weakly increasing in z. To show we first establish the following result.

Lemma B.1: $\mathbb{F}$ *is a closed set in* $\mathcal{C}$

Proof. Consider an $f' \notin \mathbb{F}$ that is the limit of a sequence $\{f_n\}, f_n \in \mathbb{F}, \forall n \in \mathbb{N}$. Then there exists an $y_1 < y$ such that $f'(x, y_1, 1) - f'(x, y_1, 2) > f'(x, y, 1) - f'(x, y, 2)$, while $f_n(x, y_1, 1) - f_n(x, y_1, 2) \leq f_n(x, y, 1) - f_n(x, y, 2)$, for every n. Define a sequence $\{s_n\}$ with $s_n = f_n(x, y_1, 1) - f_n(x, y_1, 2) - (f_n(x, y, 1) - f_n(x, y, 2))$. Then $s_n \geq 0$, $\forall n \in \mathbb{N}$. A standard result in real analysis guarantees that for any limit s of this sequence, $s_n \to s$, it holds that $s \geq 0$. Hence $f'(x, y_1, 1) - f'(x, y_1, 2) \leq f'(x, y, 1) - f'(x, y, 2)$, contradicting the premise. □

Thus, the fixed point exhibits this property as well and the optimal quit policy is a reservation-z policy given $1 - \delta - \lambda(\theta) > 0$. Since $y(A, z)$ is strictly increasing in z, the fixed point difference $M - W^U$ must also be strictly increasing in z. Furthermore, since $\lambda(\theta)$ is concave and positively valued, $\lambda'(\theta)(M - W^U) = k$ implies that job finding rate is also (weakly) increasing in z.

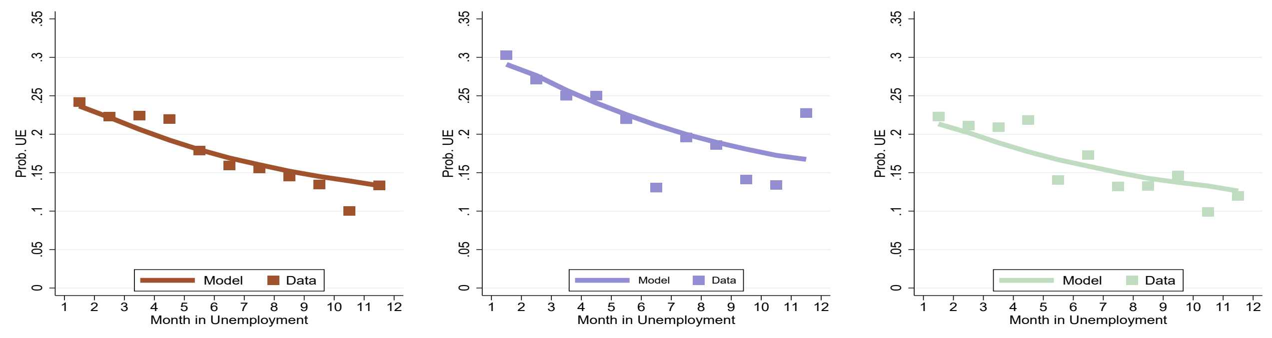
C Quantitative Analysis

This appendix is divided into three parts that complement Sections 4 and 5 of the paper. The first part provides further details of the full model calibration done in Section 4. The second part presents the calibration results from the "excess mobility model", where we analyse its ability to reproduce the long-run and cyclical patterns of several labor market variables. Here we also consider two additional excess mobility calibrations: one based on a model without human capital depreciation and the other using job spells that contain transitions between unemployment and non-participation instead of pure unemployment spells. The third part provides the details of the calibration were we shut down occupational mobility and assess the ability of a one-sector model to replicate the cyclical behaviour of unemployment and its duration distribution.

C.1 Full Model: Gross and Net Mobility

In the main text we show that the calibrated version of the full model is able to replicate well all the targeted long-run occupational mobility, job separation, job finding and unemployment patterns of the US labor market. It does so by generating within each task-based category periods of search, rest and reallocation unemployment as A, p_o and z evolve.

Here we expand on the analysis presented in Sections 4 and 5 along three dimensions. First, we further show the model's implied unemployment durations by presenting (i) the job finding rates as a function of duration (and also as a function of workers' occupational mobility status), (ii) the (incomplete) unemployment duration distribution and (iii) the relationship between occupational mobility and unemployment duration (mobility-duration profile) decomposed by the excess and net mobility. Second, we provide further details of the differences between occupational categories with respect to their relative cyclical unemployment responses, and the cyclical inflow and net flow responses that are used to estimate occupation-specific cyclical differences in the model. Third, we present the full correlation tables describing the cyclical performance of the model using the 5Q-MA smoothed and Quarterly HP-filtered measures. We also discuss the cyclicality of an alternative unemployment measure that includes entrants; show the cyclicality of the unemployment, job finding and job separation rates by age groups; and present the decomposition of search, rest and reallocation unemployment episodes for a given value of A in a comparable way to the one derived for the excess mobility mobility model discussed in Section 5.1 of the main text.



(a) All workers (b) Young workers (c) Prime-aged workers

Figure 15: Hazard Functions. Data and Model Comparison

Unemployment duration moments Figure 15 shows the aggregate and age-specific unemployment hazard functions, comparing the model to the data.[42] We observe that the model captures very well the observed duration dependence patterns, where the young exhibit a stronger decline in the job finding rate with duration than the prime-aged. Note that in our sample (and hence in the calibration) the degree of negative duration dependence in the unemployment hazard is relatively weak as we have tried to minimise the presence of unemployed workers who were in temporary layoff and/or returned to their previous employers (see the Supplementary Appendix D for a further discussion of this issue).

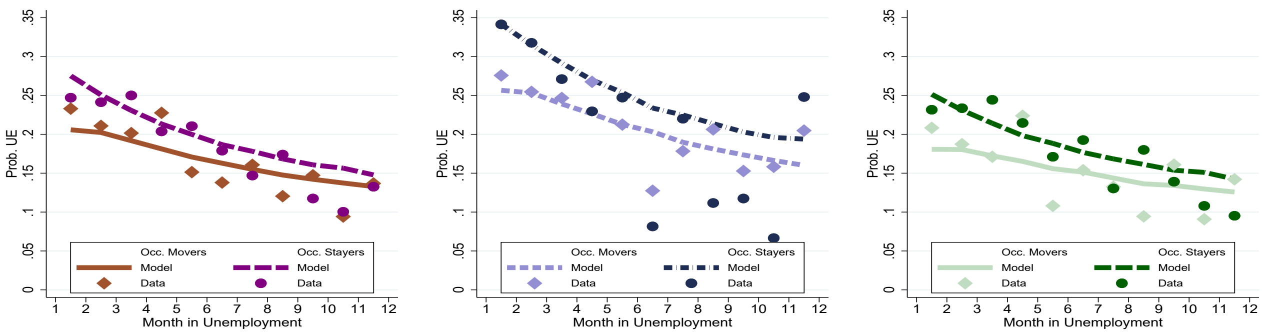
Figure 16 shows the aggregate and age-specific unemployment hazard functions separately for occupational movers and stayers. Here we observe that the model also captures well these hazards functions, separately for occupational movers and stayers, where we find both in the model and data a stronger degree of negative duration dependence among occupational stayers than occupational

[42] In the SIPP hazard functions we observe the effects of the seams present in these data. The model's estimates do not have this issue and hence are much smoother.

Table 15: Incomplete Unemployment Duration Distribution Behavior (1-18 months)

	All workers			Young workers			Prime-aged workers		
Unemp. Duration	Full Model	Excess Model	Data	Full Model	Excess Model	Data	Full Model	Excess Model	Data
1-2 m	0.43	0.42	0.43	0.53	0.52	0.47	0.40	0.39	0.41
1-4 m	0.65	0.64	0.67	0.75	0.75	0.71	0.62	0.61	0.65
5-8 m	0.20	0.21	0.20	0.17	0.17	0.19	0.22	0.22	0.21
9-12 m	0.09	0.09	0.08	0.05	0.05	0.07	0.10	0.10	0.09
13-18m	0.06	0.06	0.05	0.03	0.03	0.03	0.07	0.07	0.06

movers, particularly among young workers.



(a) All workers (b) Young workers (c) Prime-aged workers

Figure 16: Occupational Movers/Stayers Hazard Functions. Data and Model Comparison

The observed unemployment duration distribution is also well matched by the model. Table 15 shows it reproduces very well both the proportion of short and long durations spells across the distribution. Further, this fit is achieved when pooling together all workers and when separately considering young and prime-aged workers. Crucially, the fit of the duration distribution is not implied by targeting the empirical unemployment survival functions. The reported duration distribution is constructed by averaging duration distributions across quarters, while the survival functions are derived from pooling all observations. For example, the observed long-term unemployment in the pooled survival functions occur mainly in recessions and these observations get down-weighted when averaging across quarters (instead of counting each observation equally). Indeed, we show in Section 3 of this appendix that matching the survival functions does not imply also matching the duration distribution when we do not allow for occupational mobility. Instead, we show that the presence of the reallocation cutoffs are crucial to obtain a good fit in both the survival functions and the unemployment duration distributions.

In Figure 3, Section 2 of the main text we showed that both excess and net mobility increase with unemployment duration. Further, this figure showed that it is excess mobility that mainly drives the overall increase of occupational mobility with duration. Here we show that the full model calibration is also consistent with these patterns. Figure 17a shows the equivalent decomposition of the mobility-duration profile as presented in Figure 3b in the main text, using task-based categories and without considering the "management" occupation. In the model both excess and net mobility increase with

Table 16: Task-based Unemployment Duration Elasticities

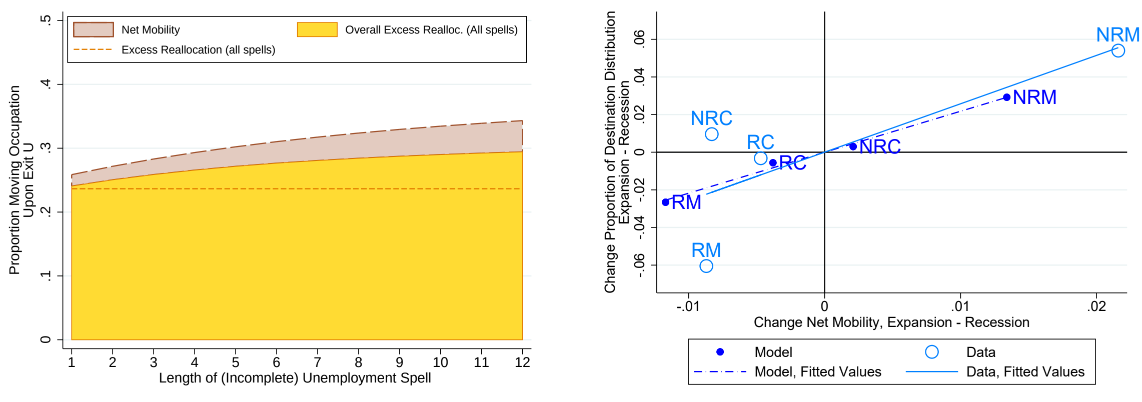
	NRC	RC	NRM	RM
$\varepsilon^{Data}_{UD_o,u}$	0.409	0.383	0.284	0.419
(s.e.)	(0.068)	(0.050)	(0.045)	(0.053)
$\varepsilon^{Model}_{UD_o,u}$	*0.390*	*0.413*	*0.342*	*0.423*
$\varepsilon^{Data}_{UD_o,u}/\varepsilon^{Data}_{UD_{avg},u}$ (targeted)	1.096	1.027	0.761	1.122
(s.e.)	(0.183)	(0.132)	(0.119)	(0.141)
$\varepsilon^{Model}_{UD_o,u}/\varepsilon^{Model}_{UD_{avg},u}$	*0.996*	*1.054*	*0.874*	*1.081*

unemployment duration. Given the countercyclicality of net mobility, the latter occurs as net mobility is more prominent in recessions where workers' unemployment durations are typically longer. Further, here we show that excess mobility is the main driver of the mobility-duration profile as in the data.

Task-based occupational categories over the cycle In the model the cyclical productivity loadings ϵ_o are the only four cyclical parameters that explicitly differ across task-based occupational categories $o \in \{NRC, RC, NRM, RM\}$. Together with the elasticity of the cross-occupation search, ν, these parameters shape the differential cyclical responses of each category o along three dimensions, summarised by 12 moments in Table 3 in the main text: (i) the cyclical responses of net mobility for each task-based category ("Net mobility o, *Recessions* and Net mobility o, *Expansions*"), (ii) the cyclical changes in the proportion of occupational movers that chooses an occupation category o ($\Delta_{exp-rec}$ (inflow $o/$all flows)), and (iii) the strength of each category's unemployment durations responses relative to economy-wide average response to the aggregate unemployment rate ($\varepsilon_{UD_o,u}/\varepsilon_{UD_{avg},u}$).

We highlight that in (iii) we target the unemployment duration elasticities for each task-based category *relative* to the economy-wide elasticity. We do this as we want to leave untargeted the amplification of aggregate unemployment. In particular, as a first step to derive these elasticities in the SIPP we regress for each task-based category the log unemployment durations of workers who lost their job in o on the log (aggregate) unemployment rate and a linear trend. Let $\varepsilon_{UD_o,u}$ for $o \in \{NRC, RC, NRM, RM\}$ denote the resulting unemployment duration elasticities with respect to aggregate unemployment. The first row of Table 16 presents these elasticities and compares them to the simulated ones in the calibration. These elasticities show that NRM occupations have a more muted cyclical response in unemployment duration than RM occupations. This differential response is also statistically significant: a Wald test on equality of the two corresponding coefficients has an associated p-value of 0.02. In the second step, we normalize each elasticity by the (occupation size-weighted) average of all four elasticities. The resulting normalized elasticities are the ones we target in the model. The last two rows of Table 16 shows these ratios (see also Table 3 in the main text), showing that model fits the data well. In particular, it shows that RM occupations are the most cyclically sensitive in terms of unemployment durations (highest value of $\varepsilon_{UD_o,u}/\varepsilon_{UD_{avg},u}$); while NRM occupations are the least cyclically sensitive (the lowest value of $\varepsilon_{UD_o,u}/\varepsilon_{UD_{avg},u}$). We observe that the models' elasticities are in line with the data. Below we show that the model is also successful in generating the untargeted aggregate unemployment amplification.

As shown in the Table 3 of the main text, the model is also consistent with the cyclical changes in net mobility as well as the cyclical changes in the inflows for each task-based category. This occurs as differences in ϵ_o translate into cyclically changing incentives for workers to leave an occupation in category o and, depending on ν, to sample z-productivities from another occupation in category o'. Figure 17b displays the relationship between these two set of moments in the model and in the data. For each task-based category, this figure shows the relationship between the cyclical changes in net mobility in the x-axis ("Net mobility o, *Expansions* - Net mobility o, *Recessions*") and the cyclical changes in inflows as a proportion of all occupational movers in the y-axis ($\Delta_{exp-rec}$ (inflow $o/$all flows)). We observe that RM occupations have the strongest cyclical response of net outflows, increasing in recessions, as well as the strongest response in the inflow proportion, also larger in recessions. In contrast, NRM occupations are the ones which experience the largest increase in net inflows in recessions and the largest increase in inflows as destination category.



(a) Decomposition of the mobility-duration profile

(b) Cyclical Shifts across Occupation Categories in Netflows vs Inflows

Figure 17: Task-Based Occupational Mobility

As Figure 17b and Table 16 show, the model captures well the co-movement along dimensions (i), (ii) and (iii). In particular, with only one set of parameters indexed by task-based occupation category, ϵ_o, the model reproduces well the average co-movement of the cyclical inflow shift with the cyclical changes in net flows. Figure 17b shows that when comparing the fitted regression lines in the data and the model both have display very similar slopes, where changing the net flow by 1% goes together with an inflow response of more than 2%.

Economy-wide cyclical outcomes In terms of the cyclical properties of the unemployment, vacancies, job finding and separation rates, Table 17 show the full set of correlations for the model and the data. The model's aggregate time series arise from the distributions of employed and unemployed workers across all labor markets, combined with agents' decisions. The top panel compares the data and model using centered 5Q-MA time series of quarterly data. The cyclical components of the (log) of these time series are obtained by using an HP filter with parameter 1600. It shows that the model

Table 17: Logged and HP-filtered Business Cycle Statistics - Full Model

Smoothed data: centred 5Q MA time series of quarterly data												
	Data (1983-2014)						**Full Model**					
	u	v	θ	s	f	$outpw$	u	v	θ	s	f	$outpw$
σ	0.14	0.11	0.25	0.10	0.09	0.01	0.14	0.05	0.17	0.07	0.10	0.01
ρ_{t-1}	0.98	0.99	0.99	0.94	0.91	0.93	0.93	0.90	0.92	0.87	0.92	0.88
	Correlation Matrix											
u	1.00	-0.92	-0.98	0.80	-0.82	-0.47	1.00	-0.61	-0.96	0.79	-0.88	-0.94
v		1.00	0.98	-0.76	0.76	0.56		1.00	0.77	-0.74	0.85	0.76
θ			1.00	-0.80	0.81	0.51			1.00	-0.83	0.95	0.96
s				1.00	-0.75	-0.39				1.00	-0.85	-0.90
f					1.00	0.27					1.00	0.93
$outpw$						1.00						1.00
Un-smoothed data												
	u	v	θ	s	f	$outpw$	u	v	θ	s	f	$outpw$
σ	0.16	0.11	0.26	0.16	0.19	0.01	0.16	0.07	0.21	0.12	0.12	0.01
ρ_{t-1}	0.85	0.96	0.94	0.58	0.42	0.75	0.86	0.47	0.82	0.34	0.72	0.76
	Correlation Matrix											
u	1.00	-0.83	-0.97	0.63	-0.58	-0.38	1.00	-0.46	-0.95	0.50	-0.78	-0.87
v		1.00	0.94	-0.71	0.57	0.45		1.00	0.72	-0.61	0.76	0.69
θ			1.00	-0.69	0.61	0.42			1.00	-0.612	0.88	0.93
s				1.00	-0.53	-0.26				1.00	-0.73	-0.76
f					1.00	0.16					1.00	0.88
$outpw$						1.00						1.00

is able to replicate the volatility and persistence of empirical time series of the unemployment, job finding and separation rates very well and generate a strong downward-sloping Beveridge curve.

To understand the reason why we present our benchmark results using a centered 5Q-MA on quarterly data, the bottom panel compares the data and model without using this smoothing procedure. The model now yields time series that for the vacancy and job separation rates are much less persistent than their data counterparts. This happens because in this case we have used a relative coarse a grid for the productivity process in the simulations, as making the productivity grid finer will make the computational time of the calibration unmanageable. This implies that the discreteness of the z^s and z^r cutoffs functions (relative to the productivity grid) then makes the vacancy and job separation rates change value too often. Using a centered 5Q-MA on quarterly data alleviates this feature without further compromising on computation time. Note, however, that this comes at the cost of slightly reducing the volatility of the vacancy rate (and labor market tightness) in the model from 0.07 to 0.05 (0.26 to 0.21), while in the data this volatility remains stable. Similarly on the data side, the job finding rates, measured in a consistent way with the model while taking into account censoring, are also somewhat noisy at quarterly frequency in the SIPP. Smoothing this time-series using the 5Q-MA helps diminish this noise.

As argued in main text (and in Supplementary Appendix G) we consider the unemployment rate of those workers who are unemployed between jobs (EUE), so that occupational mobility of these workers can be straightforwardly measured. The resulting EUE unemployment rate (EUE/(EUE+E)), under the definitions and restrictions we explained in the main text, is significantly lower than the

BLS at 3.6% (vs 6.3%), but drives much of its changes. In particular, for every one percentage point change in the BLS unemployment rate, we find that about 0.75 percentage points originate from the response of the EUE unemployment rate. This means that the relative cyclical response of the EUE unemployment rate is much stronger than the relative response of the BLS unemployment rate. Indeed, the volatility of the HP-filtered logged quarterly EUE unemployment rate is 0.16 while the corresponding BLS unemployment measure (which includes inflows from non-participation) over the same period is 0.11. For the 5Q-MA smoothed time series, the difference is from 0.14 (EUE) to 0.10 (BLS). The above also means that the focus on EUE unemployment raises the bar further to achieve sufficient amplification. Nevertheless, Table 17 shows that our model performs well.

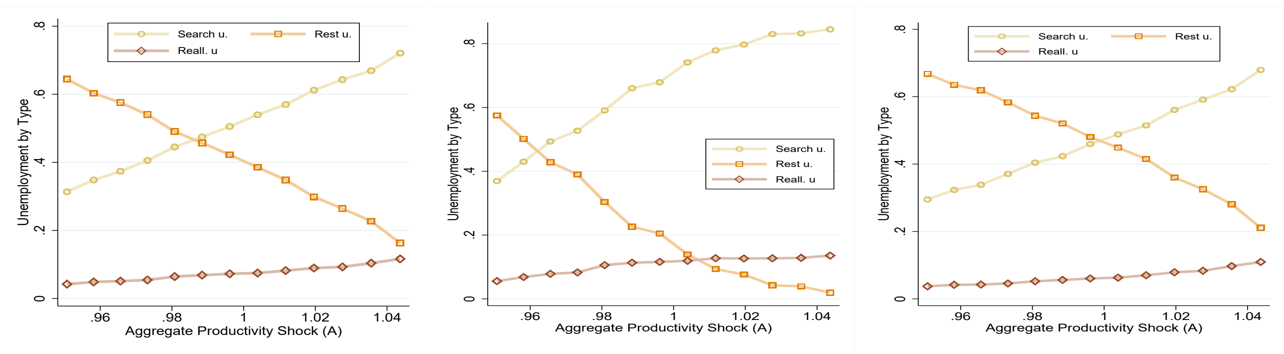
In the model we also can calculate a measure of unemployment that includes unemployment following first entry into the labor market. Relative to the BLS measure, this measure still excludes unemployment associated with workers who re-enter the labor market during their working life or who subsequently leave the labor force but not before spending time in unemployment. Including entrants in unemployment raises the average total unemployment rate to 5.2% in the model, exhibiting a lower volatility of 0.12 (5Q-MA smoothed). The latter arises as with this unemployment measure, roughly 60% of the way from the the EUE to BLS unemployment measures, its volatility gets closer to that of the BLS measure. Cross-correlation and autocorrelation statistics of this alternative unemployment measure are very similar to the EUE unemployment measure, described in Table 17.

The ability of the model to replicate the cyclical behavior of many labor market variables is down to the coexistence of episodes of search, rest and reallocation unemployment during workers' jobless spells. Figure 18 shows that when aggregating across all occupations the distribution of these types of unemployment episodes across values of A is very similar to the one generated by the excess mobility model as depicted in Figure 9c in the main text. That is, search unemployment episodes are the most common when the economy moves from mild recessions up to strong expansions. It is only as recessions get stronger that rest unemployment episodes become more common.

The middle and right panels of Figure 18 shows that among young and prime-aged workers the calibration generates similar search and rest unemployment dynamics over the business cycle as described above. This implies that we obtain high and similar cyclical volatilities for the unemployment, job finding and separation rates across age groups. In particular, the u volatilities for the young and the prime-aged are 0.139 and 0.141, the volatilities of f for young and prime-aged workers are 0.099 and 0.096; and the volatilities of s are 0.059 for young workers and 0.063 for prime-aged workers. We return to this point in the next section when presenting the calibration details of the excess mobility model.

C.2 Excess Mobility and Cyclical Unemployment

To show the importance of idiosyncratic occupation-worker (z) productivity shocks in allowing the full model to replicate the cyclical behavior of many labor market variables, we re-estimate the model by shutting down occupation-wide heterogeneity (level and business cycle loadings), effectively setting $p_{o,t} = 1$ at all t. In this case, a worker's productivity at time t in an occupation o is completely de-



(a) All workers (b) Young workers (c) Prime-aged workers

Figure 18: Unemployment decomposition - Full model

scribed by aggregate productivity A, worker-occupation match productivity z and occupation-specific human capital x_h. Workers do not (and do not want to) prefer one new occupation over another before knowing their z. Note that although we label it as the "excess mobility model", it can easily be made consistent with the observed average net flows by imposing an exogenous transition matrix that governs the probabilities with which a worker in occupation o observes a z in a different occupation o'.[43] This is in contrast to our full model, where occupational productivities $p_o, p_{o'}$ differ and change relative to one another over the cycle and in response workers change the direction of their cross-occupation search. As such the full model can be considered as the "endogenous net mobility" model, while the excess mobility model as the "exogenous net mobility" model.

C.2.1 Benchmark Excess Mobility Model

This version of the model corresponds to the excess mobility model in the main text. Except for occupation-wide productivity differences and a cross-occupation search decisions, everything else remains as described in Section 3 of the main text. We use the same functional forms as done to calibrate the full model in Section 4 of the main text. This implies that to capture economic choices and gross mobility outcomes, we now only have a set of 14 parameters to recover, where $[c, \rho_z, \sigma_z, \underline{z}_{norm}]$ governs occupational mobility due to idiosyncratic reasons (excess mobility); $[x_2, x_3, \gamma_d, \delta_L, \delta_H]$ governs differences in occupational human capital; and the remainder parameters $[k, b, \eta, \rho_A, \sigma_A]$ are shared with standard DMP calibrations. We jointly calibrate these parameters by matching the moments reported in Table 18 and Figure 19.

The excess mobility model matches very well the targeted occupational mobility moments as well as the job finding and job separation moments. The fit is comparable with the one of the full model. In particular, the excess mobility model replicates well the aggregate mobility-duration profile and the mobility-duration profiles of young and prime-aged workers. Figure 19c shows also a good fit with respect to the aggregate mobility-duration profile in expansions and recessions. Similarly, the model is able to replicate well the aggregate unemployment survival function and the survival functions of young and prime-aged workers.

This model also matches well the non-targeted moments pertaining to workers' gross occupational

[43] With a cyclically varying exogenous transition matrix, we would also be able to match the observed cyclical net flows.

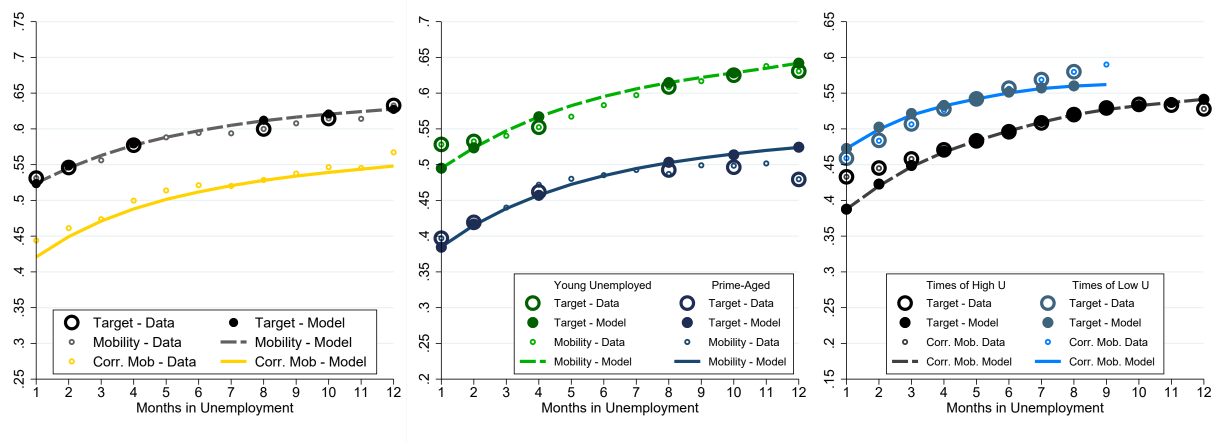
Table 18: Targeted Moments. Excess Mobility Calibration

Moments	Model	Data	Moments	Model	Data	Moments	Model	Data
Aggregate Productivity			**U Survival w. Age**			**Returns to Human Capital**		
outpw	1.005	1.000	Young 2 months	0.713	0.697	5 years (OLS)	0.150	0.154
ρ_{outpw}	0.760	0.753	Young 4 months	0.387	0.381	10 years (OLS)	0.230	0.232
σ_{outpw}	0.0094	0.0094	Young 8 months	0.146	0.156			
Aggregate Matching Function			Young 12 months	0.069	0.073	**Empirical Separation moments**		
$\hat{\eta}$	0.506	0.500	Young 16 months	0.037	0.038	rel. sep rate young/prime	2.146	2.004
Unemployment Rate			Young 20 months	0.020	0.020	prob (u within 3yrs for empl.)	0.148	0.124
u	0.0355	0.0355				rel sep rate recent hire/all	5.221	4.945
U. Survival all workers			Prime 2 months	0.783	0.777			
2 months	0.763	0.758	Prime 4 months	0.506	0.485	**Cyclical Mobility-Duration Profile Shift**		
4 months	0.472	0.457	Prime 8 months	0.251	0.234	Times Low U. - 1 month	0.473	0.459
8 months	0.221	0.208	Prime 12 months	0.142	0.137	Times Low U. - 2 months	0.503	0.484
12 months	0.120	0.120	Prime 16 months	0.086	0.090	Times Low U. - 3 months	0.522	0.507
16 months	0.071	0.076	Prime 20 months	0.055	0.061	Times Low U. - 4 months	0.533	0.528
20 months	0.045	0.048				Times Low U. - 5 months	0.542	0.542
						Times Low U. - 6 months	0.551	0.557
Occ. Mobility-Duration Profile All			**Occ. Mobility-Duration Profile w. Age**			Times Low U. - 7 months	0.557	0.569
1 month	0.523	0.531	Young 2 months	0.613	0.608	Times Low U. - 8 months	0.560	0.580
2 months	0.548	0.546	Young 4 months	0.646	0.613			
4 months	0.579	0.577	Young 8 months	0.685	0.669	Times High U. -1 month	0.388	0.433
8 months	0.612	0.600	Young 10 months	0.695	0.679	Times High U. -2 months	0.423	0.445
10 months	0.621	0.615	Young 12 months	0.706	0.725	Times High U. -3 months	0.449	0.458
12 months	0.627	0.633				Times High U. -4 months	0.469	0.471
			Prime 2 months	0.520	0.513	Times High U. -5 months	0.484	0.483
			Prime 4 months	0.553	0.556	Times High U. -6 months	0.497	0.496
			Prime 8 months	0.591	0.568	Times High U. -7 months	0.511	0.509
			Prime 10 months	0.599	0.577	Times High U. -8 months	0.520	0.520
			Prime 12 months	0.606	0.565	Times High U. -9 months	0.529	0.531
						Times High U. -10 months	0.532	0.536
						Times High U. -11 months	0.537	0.535
						Times High U. -12 months	0.541	0.528

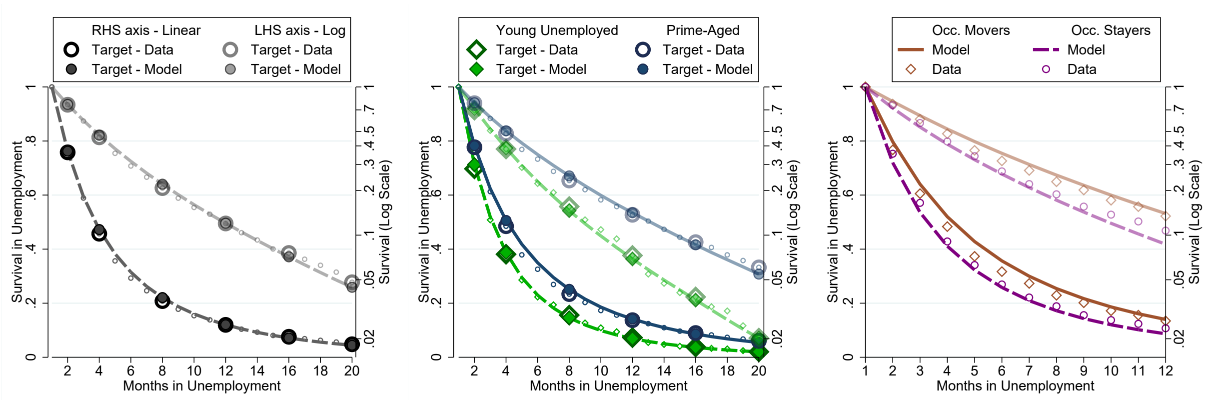
mobility and job finding hazards discussed in the previous section. For example, Table 15 shows that the excess mobility model is able to reproduce the observed unemployment duration distribution for all workers and by age groups. The fit of other untargeted moments is not shown here to save space, but available upon request. The estimated parameter values in this calibration are also very similar to the ones obtained in the full model. These are $c = 7.549$, $k = 125.733$, $b = 0.843$, $\eta = 0.241$, $\delta_L = 0.0034$, $\delta_H = 0.0004$, $\underline{z}_{corr} = 0.349$, $\rho_A = 0.998$, $\sigma_A = 0.00198$, $\rho_z = 0.998$, $\sigma_z = 0.00707$, $x_2 = 1.181$, $x_3 = 1.474$ and $\gamma_h = 0.0039$.

The first key insight from this exercise is that to match the targeted gross occupational mobility, job finding and job separation moments one does not need endogenous net mobility. Instead this calibration highlights the role of worker-occupation idiosyncratic productivity shocks and human capital accumulation in fitting all of the above patterns.

As shown in the main text, the excess mobility calibration is also able to fit a wide range of cyclical features of the labor market. The left panel of Table 20 (below) shows the time series properties of the unemployment, vacancy, job finding and job separation rates and of labor market tightness as well as the full set of correlations between them, obtained from the excess mobility calibration. Here we find that the cyclical implications of the excess mobility model are very similar to that of the full model. These results highlight the second key insight from this exercise: endogenous net mobility does not play an important role in making the model replicate the aforementioned cyclical labor market features. As shown in the main text, in Table 6, the same conclusion holds when evaluating the role of



(a) Mobility-Duration Profile (All) (b) Mobility-Duration Young-Prime (c) Cyclical Mobility-Duration Profile



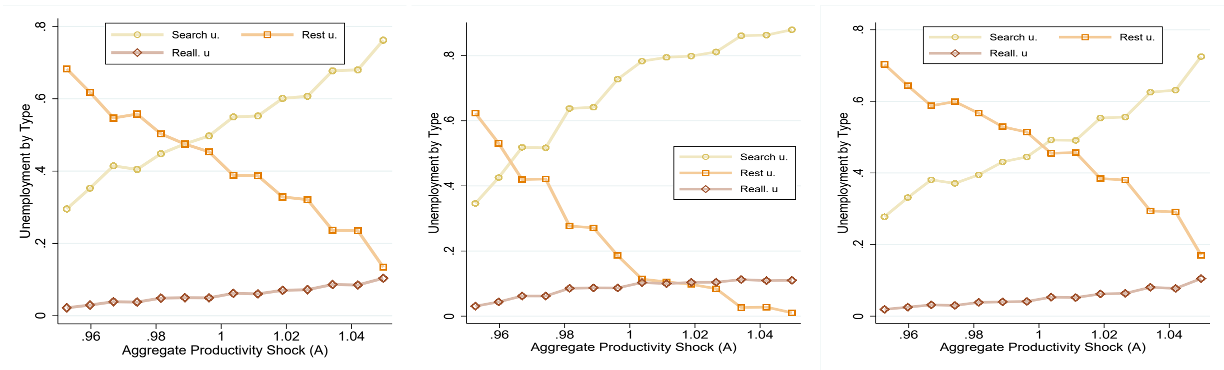
(d) Survival Profile (All) (e) Survival Profile Young-Prime (f) Survival Profile Movers-Stayers

Figure 19: Targeted Moments. Data and Model Comparison

endogenous net mobility in making the model replicate the cyclical behaviour of the unemployment duration distribution.

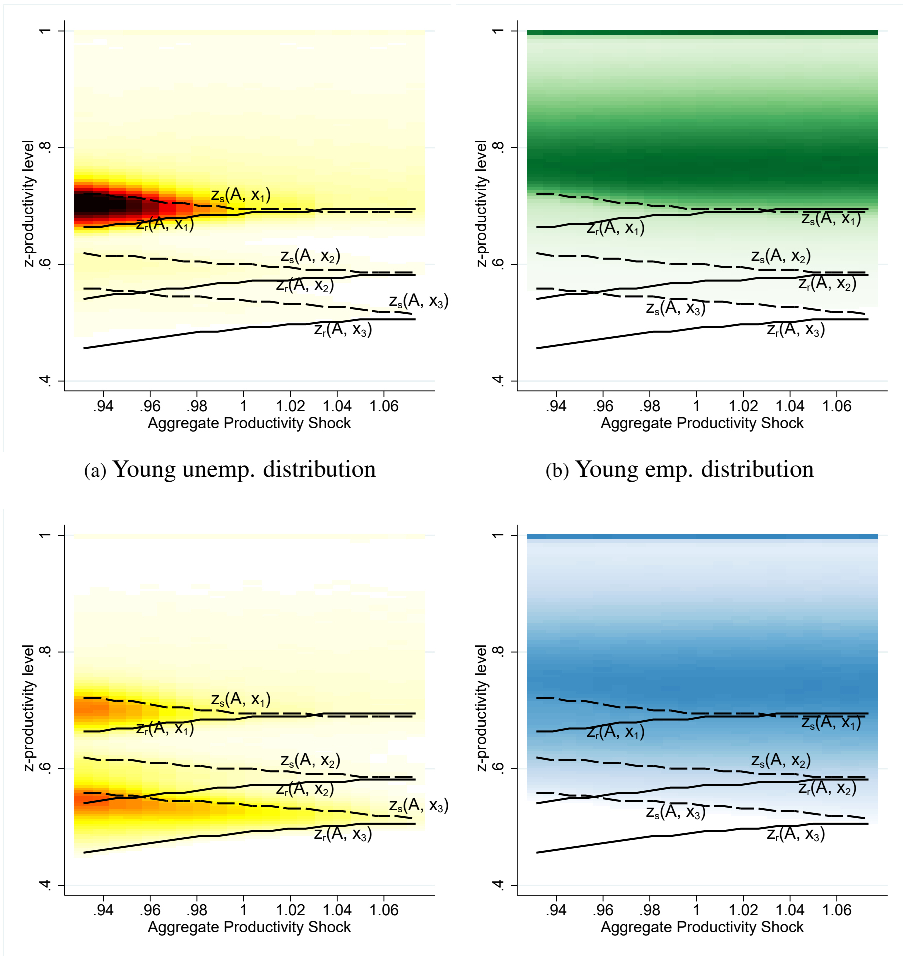
The main reason why the excess mobility model is able to replicate all these cyclical features is because the importance of search, rest and reallocation unemployment episodes during a jobless spell is driven by the interaction between the aggregate shock and the worker-occupation match z-productivity. The difference between the z^s and z^r cutoffs creates an area of inaction that widens during recessions "trapping" workers for a longer time in rest unemployment episodes. As the economy recovers the difference between these cutoffs narrows and the area of inaction shrinks allowing workers to escape by crossing both the z^s and z^r cutoffs. These features then yield the procyclicality of gross (and excess) occupational mobility and the countercyclicality of job separations. Hence, the decomposition in Figure 20a looks very similar to Figure 18a for the full model.

Age patterns Figure 20 shows that the above dynamics not only happen when pooling all workers together but for each age group, as in the full model. Figure 21 shows these age group dynamics more clearly by depicting the distribution of unemployed and employed workers among young and prime-aged workers. It shows that during recessions unemployment among young workers is concentrated both slightly above $z^s(., x_1)$ and between $z^s(., x_1)$ and $z^r(., x_1)$. During expansions, however, unemployment is located above the $z^s(., x_1)$ cutoff. In the case of prime-aged workers, the concentration of



(a) All workers (b) Young workers (c) Prime-aged workers

Figure 20: Unemployment decomposition - Excess mobility model



(a) Young unemp. distribution (b) Young emp. distribution

(c) Prime-aged unemp. distribution (d) Prime-aged emp. distribution

Figure 21: Unemployment decomposition and aggregate productivity by age groups

unemployment during recessions and expansions occurs mostly above $z^s(., x_3)$ and between $z^s(., x_3)$ and $z^r(., x_3)$, but also between the $z^s(., x_1)$ and $z^r(., x_1)$ cutoffs. This difference implies that during expansion episodes of rest unemployment are still prevalent among prime-aged workers, while for young workers these episodes basically disappear as shown in Figure 20 and consistent with a lower occupational mobility rate among prime-aged workers.

As in the full model, the excess mobility calibration obtains a similar cyclicality for the unemployment, job finding and separations rates across age-groups. In both models this occurs because the estimated z-productivity process places enough workers on the z^s cutoffs across the respective human capital levels. Figure 21 shows that, as is the case for low human capital workers, many high human capital workers enjoy good z-productivities, but for a number of them their z-productivity has drifted down and has positioned them close to $z^s(.,x_3)$. Some of these high human capital workers will subsequently leave the occupation, but over time the stock of workers close to $z^s(.,x_3)$ will be replenished by those workers who currently have high z-productivities but will suffer bad z-realizations in near future. As $z^s(.,x_3) < z^s(.,x_1)$ we observe that the average level of separations is lower for high human capital workers, but this nevertheless does not preclude the similarity in the aforementioned cyclical responsiveness.

Given that it is clear the excess mobility model is able to replicate on its own many critical features of the full model, in what follows we use it to perform two key exercises. The first one highlights the effect of human capital depreciation in attenuating the cyclical properties of the above labor market variables and motivates our use of the cyclical shift of the mobility-duration profile as a target. The second exercise investigates the quantitative implications of our model when considering that a worker's varying job finding prospects (due to the stochastic nature of the z-productivity process) during a jobless spell can be linked to observed transitions between the states of unemployment and non-participation (or marginally attached to the labour force) as defined in the SIPP. For this exercise we recompute all of the relevant empirical targets using non-employment spells that contain a mix of periods of unemployment and non-participation. We refer to this last exercises in the Conclusions of the main text.

C.2.2 The Importance of Human Capital Depreciation

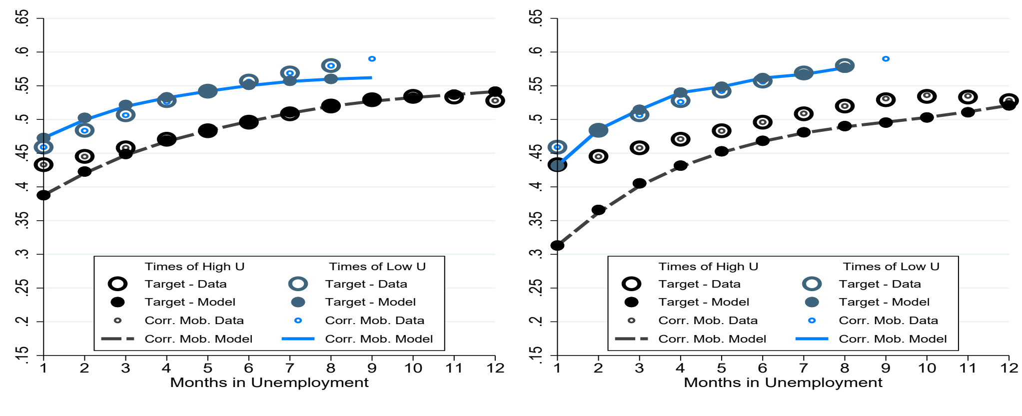
To estimate the full and the excess mobility models we used the mobility-duration profiles at different durations during recessions and expansions. These patterns informed us about the rate of occupational human capital depreciation during spells of unemployment. In the main text, we argued that these profiles were crucial in helping us identify the depreciation parameter, γ_h. The reason for the latter is that a model which did not incorporate human capital depreciation will generate very similar long-run moments as a model which did incorporate depreciation, but generate different cyclical predictions. To show this, we now present the estimation results from the excess mobility model without human capital depreciation. We target the same *long-run* moments as in the calibration described above, but do not target the cyclical behaviour of the mobility-duration profile.

Table 19 shows that the fit of the model is very good, similar to the models which incorporates human capital depreciation. Although not shown here, it also does well in matching the same untargeted long-run moments described above. The estimated parameter values are also similar with $c = 9.853$, $k = 152.073$, $b = 0.820$, $\eta = 0.181$, $\delta_L = 0.0025$, $\delta_H = 0.0008$, $\underline{z}_{corr} = 0.407$, $\rho_A = 0.997$, $\sigma_A = 0.0019$, $\rho_z = 0.999$, $\sigma_z = 0.0053$, $x_2 = 1.158$ and $x_3 = 1.491$. Further, this calibration finds that periods of search, rest and reallocation unemployment can arise during a worker's jobless spell

Table 19: Targeted Moments. No Occupational Human Capital Depreciation

Moments	Model	Data	Moments	Model	Data	Moments	Model	Data
Aggregate Productivity			**U Survival w. Age**			**Returns to Human Capital**		
outpw	1.001	1.000	Young 2 months	0.721	0.697	5 years (OLS)	0.148	0.154
ρ_{outpw}	0.776	0.753	Young 4 months	0.406	0.381	10 years (OLS)	0.246	0.232
σ_{outpw}	0.0093	0.0094	Young 8 months	0.161	0.156			
Aggregate Matching Function			Young 12 months	0.075	0.073	**Empirical Separation moments**		
$\hat{\eta}$	0.503	0.500	Young 16 months	0.039	0.038	rel. sep rate young/prime	1.944	2.004
Unemployment Rate			Young 20 months	0.021	0.020	prob (u within 3yrs for empl.)	0.141	0.124
u	0.0358	0.0355				rel sep rate recent hire/all	6.311	4.945
U. Survival all workers			Prime 2 months	0.749	0.777			
2 months	0.744	0.758	Prime 4 months	0.480	0.485			
4 months	0.460	0.457	Prime 8 months	0.246	0.234			
8 months	0.223	0.208	Prime 12 months	0.143	0.137			
12 months	0.124	0.120	Prime 16 months	0.089	0.090			
16 months	0.075	0.076	Prime 20 months	0.057	0.061			
20 months	0.048	0.048						
Occ. Mobility-Duration Profile All			**Occ. Mobility-Duration Profile Young**			**Occ. Mobility-Duration Profile Prime**		
1 month	0.481	0.532	Young 2 months	0.581	0.608	Prime 2 months	0.496	0.513
2 months	0.520	0.546	Young 4 months	0.632	0.613	Prime 4 months	0.542	0.556
4 months	0.567	0.576	Young 8 months	0.688	0.669	Prime 8 months	0.584	0.568
8 months	0.612	0.605	Young 10 months	0.696	0.679	Prime 10 months	0.594	0.577
10 months	0.619	0.622	Young 12 months	0.709	0.725	Prime 12 months	0.599	0.565
12 months	0.627	0.639						

across all levels of occupational human capital; i.e $z^s > z^r$ for all A and x_h.



(a) Model with HC Dep. (b) Model without HC Dep.

Figure 22: Cyclical Mobility-Duration Profile

Figure 22 shows the first key difference between the excess mobility model with and without occupational human capital depreciation. We plot the mobility-duration profile in times of expansions and recessions (low and high unemployment, respectively), where Figure 22a shows the mobility-duration profiles from the model with human capital depreciation and Figure 22b shows the ones for the model without human capital depreciation. The latter finds that the lack of human capital depreciation does not allow the model to match the mobility-duration profile at low durations during expansions and completely misses the profile at all durations during recessions. Is precisely this lack of fit that motivated us to add the cyclical patterns of the mobility-duration profile as targets in order to help identify the rate of human capital depreciation.

Table 20 shows the second key difference. The model without depreciation generates a larger

Table 20: Logged and HP-filtered Business Cycle Statistics

	Excess Mobility Model with HC dep.						**Excess Mobility Model with No HC dep.**					
	u	v	θ	s	f	$outpw$	u	v	θ	s	f	$outpw$
σ	0.14	0.05	0.18	0.07	0.10	0.01	0.20	0.06	0.25	0.10	0.14	0.01
ρ_{t-1}	0.95	0.89	0.94	0.88	0.93	0.94	0.94	0.87	0.94	0.88	0.94	0.94
	Correlation Matrix						**Correlation Matrix**					
u	1.00	-0.63	-0.97	0.78	-0.88	-0.94	1.00	-0.62	-0.97	0.78	-0.89	-0.92
v		1.00	0.80	-0.68	0.85	0.77		1.00	0.76	-0.61	0.77	0.72
θ			1.00	-0.81	0.95	0.96			1.00	-0.78	0.94	0.93
s				1.00	-0.82	-0.87				1.00	-0.81	-0.84
f					1.00	0.93					1.00	0.89
$outpw$						1.00						1.00

Note: Each model's aggregate time series arise from the distributions of employed and unemployed workers across all labor markets, combined with agents' decisions. Times series are centered 5Q-MA series of quarterly data to smooth out the discreteness in the relatively flat cutoffs (relative to the grid). The cyclical components of the (log) of these time series are obtained by using an HP filter with parameter 1600.

amount of cyclical volatility in the aggregate unemployment, job finding and job separation rates in relation to the model with human capital depreciation. Relative to the data, Table 17 shows an overshooting in the volatilities of the unemployment and job finding rates. To understand why this is the case and why it misses on the cyclical shift of the mobility-duration profile, Figure 23 presents the distribution of search, rest and reallocation unemployment episodes for each level of A. This figure shows that the calibration without occupational human capital depreciation also has the property that rest unemployment is the more prevalent episode during recessions while search unemployment is the more prevalent episode during expansions among all workers and by age groups.

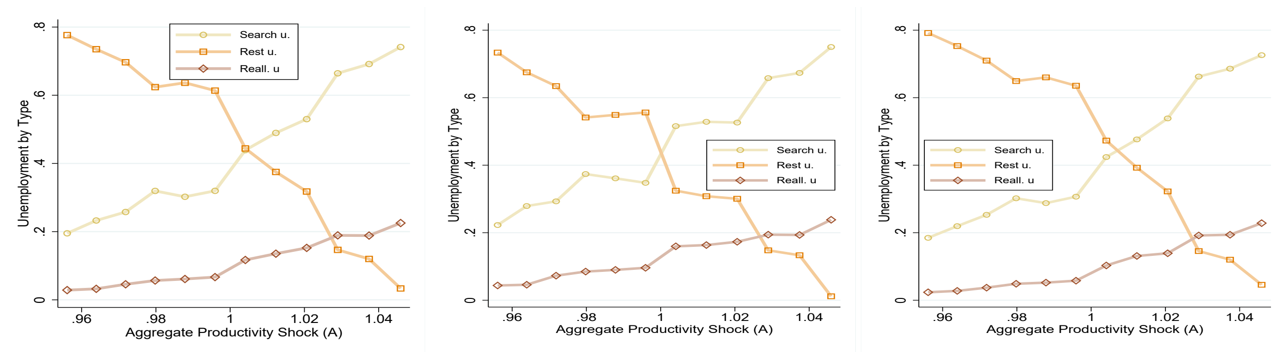


Figure 23: Unemployment decomposition - No occupational human capital depreciation

This mechanism, however, becomes more powerful when we do not include human capital depreciation. Figure 23 shows this through a sharper drop in the proportion of rest unemployment and a sharper rise in the proportions of search and reallocation unemployment as the economy improves. Human capital depreciation attenuates these effects. In particular, job separations become somewhat less countercyclical because now workers take into account that if they decide to separate they will face the prospect of human capital loss and hence lower job finding rates. At the same time rest unemployed workers with human capital levels x_2 and x_3 now face a lower expected opportunity cost of mobility (even if their z would improve, a depreciation shock might trigger a reallocation anyway), leading to a lower proportion of rest unemployment episodes for all A and crucially to a significantly less procyclical job finding and unemployment rates.

C.2.3 The Unemployed and Marginally Attached

In the previous calibrations we built the analysis based on the interpretation that, although a worker who is currently in a rest unemployment episode cannot find a job, he would want to search for jobs (as opposed to stay idle at home) because he still faces a positive expected job finding probability in the near future. Episodes of rest unemployment, however, could conceptually be extended to incorporate marginally attached workers. To investigate the latter we expand our analysis to capture more broadly the occupational mobility decisions of the unemployed and marginally attached in shaping the cyclicality of aggregate unemployment.

We do this by re-estimating the excess mobility model, recomputing the targets using non-employment spells in which workers transition between unemployment and non-participations as labelled in the SIPP. In particular, we consider non-employment spells with at least one period of unemployment, which we label ‘NUN’ spells. To avoid maternity and related issues in non-participation we restrict the focus to men. We show that when considering non-participation periods the model still reproduces the observed cyclical amplification in the non-employment rate of those unemployed and marginally attached workers.[44]

Table 21: Targeted Moments. NUN spells

Moments	Model	Data	Moments	Model	Data	Moments	Model	Data
Aggregate Productivity			**U Survival w. Age**			**Returns to Human Capital**		
outpw	1.005	1.000	Young 2 months	0.783	0.797	5 years (OLS)	0.156	0.154
ρ_{outpw}	0.781	0.753	Young 4 months	0.527	0.515	10 years (OLS)	0.257	0.232
σ_{outpw}	0.0093	0.0094	Young 8 months	0.273	0.290			
Aggregate Matching Function			Young 12 months	0.159	0.172	**Empirical Separation moments**		
$\hat{\eta}$	0.503	0.500	Young 16 months	0.097	0.116	rel. sep rate young/prime	2.263	2.004
NUN nonemployment Rate			Young 20 months	0.062	0.080	prob (u within 3yrs for empl.)	0.161	0.124
NUN/(NUN+E)	0.053	0.052				rel sep rate recent hire/all	4.848	4.945
U. Survival all workers			Prime 2 months	0.832	0.853			
2 months	0.818	0.836	Prime 4 months	0.608	0.586	**Cyclical Mobility-Duration Profile Shift**		
4 months	0.585	0.570	Prime 8 months	0.357	0.334	Times Low U. - 1 month	0.464	0.454
8 months	0.334	0.326	Prime 12 months	0.227	0.216	Times Low U. - 2 months	0.484	0.474
12 months	0.208	0.213	Prime 16 months	0.150	0.157	Times Low U. - 3 months	0.497	0.493
16 months	0.135	0.153	Prime 20 months	0.103	0.115	Times Low U. - 4 months	0.509	0.522
20 months	0.092	0.117				Times Low U. - 5 months	0.521	0.545
						Times Low U. - 6 months	0.531	0.557
Occ. Mobility-Duration Profile All			**Occ. Mobility-Duration Profile w. Age**			Times Low U. - 7 months	0.545	0.546
1 month	0.522	0.537	Young 2 months	0.593	0.593	Times Low U. - 8 months	0.552	0.544
2 months	0.543	0.551	Young 4 months	0.618	0.615			
4 months	0.572	0.590	Young 8 months	0.652	0.658	Times High U. -1 month	0.416	0.441
8 months	0.613	0.623	Young 10 months	0.665	0.678	Times High U. -2 months	0.445	0.458
10 months	0.629	0.650	Young 12 months	0.675	0.719	Times High U. -3 months	0.466	0.477
12 months	0.640	0.677				Times High U. -4 months	0.486	0.508
			Prime 2 months	0.522	0.520	Times High U. -5 months	0.504	0.512
			Prime 4 months	0.553	0.570	Times High U. -6 months	0.517	0.531
			Prime 8 months	0.595	0.590	Times High U. -7 months	0.532	0.536
			Prime 10 months	0.613	0.613	Times High U. -8 months	0.544	0.555
			Prime 12 months	0.625	0.619	Times High U. -9 months	0.556	0.578
						Times High U. -10 months	0.564	0.608
						Times High U. -11 months	0.572	0.605
						Times High U. -12 months	0.582	0.639

[44] Note that here we also focus on spells of at least one month, and workers who say that they “without a job”, mirroring these sample restrictions for unemployment spells.

Table 21 shows the targets and the fit of this estimation. As documented in more detail in the Supplementary Appendix and in Section 2 of the main text, the mobility-duration profile including NUN spells does not differ much from the profile of only the unemployed. The survival probability in NUN spells, however, shifts up significantly at longer durations, compared to the corresponding patterns for unemployment spells, both for all workers and across age groups. For example, pooling the entire sample, around 10% of NUN spells last 20 months or more (relative to less than 5% for unemployment spells); while even for young workers around 8% of NUN spells last more 20 months or more (relative to about 2% for unemployment spells). Including the marginally attached also implies a higher jobless rate. Nevertheless, the model can capture these features well, as it does for the other moments, including the cyclical shift of the mobility-duration profile.

The estimated parameter values are also broadly similar to the ones in the previous versions of the excess mobility model, changing in expected directions. In this case we obtain that $c = 10.822$, $k = 3.161$, $b = 0.804$, $\eta = 0.524$, $\delta_L = 0.0046$, $\delta_H = 0.0015$, $\underline{z}_{corr} = 0.428$, $\rho_A = 0.998$, $\sigma_A = 0.0020$, $\rho_z = 0.997$, $\sigma_z = 0.0134$, $x_2 = 1.146$, $x_3 = 1.712$ and $\gamma_d = 0.0084$. Note that the z-productivity process is now somewhat more volatile, but the higher reallocation cost implies that the area of inaction between the separation and reallocation cutoffs is (in relative terms) also larger. The latter leaves more scope for workers to get "trapped" for longer periods in rest unemployment episodes, thus creating an increase in the survival functions across all, young and prime-aged workers as observed in the data. Further, although k is estimated to have a much smaller value, the cost of posting a vacancy in this version of the model is actually higher than in our benchmark calibration at 0.986 of weekly output. We also estimate the elasticity of the matching function in each submarket to be about twice as big as the one in the benchmark calibration. These differences, however, do not affect our main conclusions.

Table 22: Logged and HP-filtered Business Cycle Statistics

	Data (1983-2014) - NUN spells						**Excess Mobility Model - NUN spells**					
	u	v	θ	s	f	$outpw$	u	v	θ	s	f	$outpw$
σ	0.09	0.11	0.19	0.09	0.07	0.01	0.09	0.04	0.11	0.05	0.08	0.01
ρ_{t-1}	0.98	0.99	0.99	0.94	0.91	0.93	0.95	0.81	0.93	0.84	0.92	0.94
	Correlation Matrix						**Correlation Matrix**					
u	1.00	-0.91	-0.97	0.74	-0.96	-0.40	1.00	-0.40	-0.95	0.63	-0.80	-0.87
v		1.00	0.98	-0.76	0.91	0.56		1.00	0.66	-0.37	0.61	0.53
θ			1.00	-0.77	0.95	0.48			1.00	-0.64	0.86	0.89
s				1.00	-0.84	-0.39				1.00	-0.59	-0.75
f					1.00	0.36					1.00	0.83
$outpw$						1.00						1.00

Note: Each model's aggregate time series arise from the distributions of employed and unemployed workers across all labor markets, combined with agents' decisions. Times series are centered 5Q-MA series of quarterly data to smooth out the discreteness in the relatively flat cutoffs (relative to the grid). The cyclical components of the (log) of these time series are obtained by using an HP filter with parameter 1600.

Table 22 shows the main takeaway of this exercise. The model remains able to generate cyclical movements of the non-employment, job finding and job separation rates as well as a relatively strong Beveridge curve. In particular, the cyclical volatilities of the non-employment and job finding rates

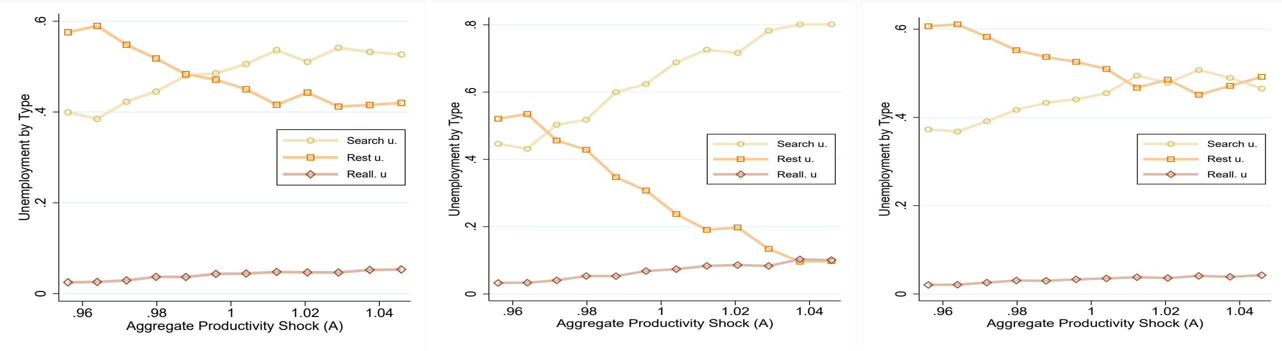


Figure 24: Unemployment decomposition - NUN spells

are the same as in the data. As in the previous estimations, here we also find that the reason for the amplification of the non-employment rate is that the model generates period of search, rest and reallocation unemployment, whose relative importance changes over the cycle. Figure 24 shows that episodes of rest unemployment are the more prevalent type during recessions while episodes of search unemployment are the more prevalent type during expansion.

Note that including marginally attached workers in our analysis increases the overall importance of rest unemployment in normal times. This is consistent with the fact that in these times the non-employment rate is higher and the associated job finding rate lower, compared to our benchmark unemployment and job finding rates measures. Further, this version of the model still needs to accommodate short-term outflows as before and does so mostly through search unemployment episodes. As a result, the proportion of rest unemployment decreases at a slower rate with A. Even at the highest aggregate productivity levels rest unemployment is very prevalent, representing about 40% of all episodes during a non-employment spell, with a large role for prime-aged workers.

Overall we find that a version of the excess mobility model that considers NUN spells exhibits a higher non-employment rate but a lower cyclicality (logging and HP-filtering the data) than in our benchmark model. This is consistent with the data, where we observe a lower cyclicality among the non-employment than among the unemployment.

C.3 The Importance of Occupational Mobility

To demonstrate that it is crucial for the long-run and cyclical performance of our model to allow for occupational mobility, such that workers in episodes of rest unemployment face both the separation and reallocation cutoffs, we re-estimate the model by shutting down the possibility of occupational mobility. This is done by exogenously setting c to a prohibiting level. We present two calibrations with no occupational mobility. Model I targets the same moments as in the full model with the exception of those pertaining to occupational mobility itself. We evaluate its fit and the implied cyclical patterns, finding that this model replicates well the long-tun targets but misses on the cyclical patterns of the unemployment, job finding and separation rates.[45] To gain further insights into the working of the no

[45] To make the estimation of Model I as comparable as possible with the previous ones, we continue targeting the returns to occupational mobility to inform the human capital levels, x_2 and x_3. Under no occupational mobility, human capital could also be interpreted as general and not occupation specific, depending on the aim of the exercise. For this exercise it would be more appropriate to target the returns to general experience. However, a comparison between the OLS returns

Table 23: Targeted Moments. No Occupational Mobility I

Moments	Model	Data	Moments	Model	Data	Moments	Model	Data
Aggregate Productivity			**U Survival w. Age**			**Returns to Human Capital**		
outpw	1.005	1.000	Young 2 months	0.678	0.697	5 years (OLS)	0.151	0.154
ρ_{outpw}	0.787	0.753	Young 4 months	0.347	0.381	10 years (OLS)	0.240	0.232
σ_{outpw}	0.0092	0.0094	Young 8 months	0.133	0.156			
Aggregate Matching Function			Young 12 months	0.069	0.073	**Empirical Separation moments**		
$\hat{\eta}$	0.390	0.500	Young 16 months	0.041	0.038	rel. sep rate young/prime	2.125	2.004
Unemployment Rate			Young 20 months	0.026	0.020	prob u within 3yrs for emp.	0.181	0.124
u	0.0358	0.0355				rel sep rate recent hire/all	3.023	4.945
U. Survival all workers			Prime 2 months	0.758	0.777			
2 months	0.735	0.758	Prime 4 months	0.481	0.485			
4 months	0.442	0.457	Prime 8 months	0.246	0.234			
8 months	0.213	0.208	Prime 12 months	0.144	0.137			
12 months	0.123	0.120	Prime 16 months	0.089	0.090			
16 months	0.075	0.076	Prime 20 months	0.057	0.061			
20 months	0.048	0.048						

occupational mobility case, Model 2 is chosen to achieve a higher cyclical volatility in the aggregate but is more permissive of deviations from the targets.

Model I Table 23 presents all the targeted moments and shows that the fit is largely comparable along nearly all corresponding dimensions to the full and excess mobility models. The parameter estimates in this calibration remain largely sensible. In this case we obtain that $k = 195.58$, $b = 0.608$, $\eta = 0.290$, $\delta_L = 0.0092$, $\delta_H = 0.0014$, $\underline{z}_{corr} = 0.258$, $\rho_A = 0.9983$, $\sigma_A = 0.0021$, $\rho_z = 0.9923$, $\sigma_z = 0.0300$, $x_2 = 1.184$, $x_3 = 1.387$ and $\gamma_d = 0.00244$. Note, however, that we now have a higher role for search frictions as the model estimates a higher value k. Further, the z-productivity process is now less persistent and exhibits a much larger variance in the stationary distribution, generating a perhaps too large Mm ratio of 2.28.[46] The higher volatility of the z process may have created some difficulty to hit the ratio of separations of recently hired workers to all workers, but overall the long-run moments are matched well. This is in contrast to its cyclical patterns.

Table 24 under "No Occupational Mobility - Model I" shows that a model that does not allow for occupational mobility, but reproduces well almost all the moments in Table 23 cannot generate enough cyclical volatility on all the relevant labor market variables. The unemployment, job finding and separation rates exhibit below half the volatility relative to their counterparts in the models with occupational mobility. It also generates a much weaker negative correlation between unemployment and vacancies.

Moreover, Model I is not able to reproduce the observed average quarterly unemployment duration distribution at short and long durations, nor does it capture the cyclical behavior of this distribution. While Model I and the occupational mobility models replicate the same unemployment survival functions, they generate different incomplete duration distributions. This occurs because the survival

to general experience and the OLS returns to occupational human capital estimated by Kambourov and Manovskii (2009) from the PSID (see their Table 3 comparing columns 1 and 3 or 6 and 8), suggests that this bias should be moderate. Using their estimates, the 5 year returns to general experience is about 0.19, while the 10 years returns is about 0.38.

[46] In this context the z-productivity process can be interpreted as an idiosyncratic productivity shock affecting a worker's overall productivity, rather than a worker's idiosyncratic productivity within an occupation.

Table 24: Logged and HP-filtered Business Cycle Statistics

	No Occupational Mobility - Model I						No Occupational Mobility - Model II					
	u	v	θ	s	f	$outpw$	u	v	θ	s	f	$outpw$
σ	0.04	0.02	0.06	0.03	0.03	0.01	0.10	0.03	0.12	0.08	0.05	0.01
ρ_{t-1}	0.94	0.84	0.93	0.85	0.86	0.94	0.95	0.83	0.94	0.90	0.89	0.94
	Correlation Matrix						**Correlation Matrix**					
u	1.00	-0.32	-0.92	0.72	-0.72	-0.85	1.000	-0.54	-0.98	0.84	-0.77	-0.97
v		1.000	0.67	-0.18	0.48	0.51		1.00	0.69	-0.55	0.69	0.61
θ			1.00	-0.63	0.77	0.88			1.00	-0.85	0.82	0.97
s				1.00	-0.48	-0.73				1.00	-0.84	-0.90
f					1.00	0.65					1.00	0.83
$outpw$						1.00						1.00

Note: Each model's aggregate time series arise from the distributions of employed and unemployed workers across all labor markets, combined with agents' decisions. Times series are centered 5Q-MA series of quarterly data to smooth out the discreteness in the relatively flat cutoffs (relative to the grid). The cyclical components of the (log) of these time series are obtained by using an HP filter with parameter 1600.

Table 25: Incomplete Unemployment Duration Distribution Behavior

Panel A: Incomplete Unemployment Distribution (1-18 months)

	All workers				Young workers				Prime-aged workers			
Unemp. Duration	Occ Model	No Occ. Model I	No Occ. Model II	Data	Occ. Model	No Occ. Model I	No Occ. Model II	Data	Occ. Model	No Occ. Model I	No Occ. Model II	Data
1-2 m	0.43	0.36	0.35	0.43	0.53	0.44	0.35	0.47	0.40	0.34	0.35	0.41
1-4 m	0.65	0.57	0.55	0.67	0.75	0.65	0.54	0.71	0.62	0.54	0.55	0.65
5-8 m	0.20	0.22	0.22	0.20	0.17	0.19	0.22	0.19	0.22	0.23	0.23	0.21
9-12 m	0.09	0.12	0.12	0.08	0.05	0.09	0.13	0.07	0.10	0.13	0.12	0.09
13-18m	0.06	0.09	0.10	0.05	0.03	0.07	0.11	0.03	0.07	0.10	0.10	0.06

Panel B: Cyclical Changes of the Incomplete Unemployment Distribution (1-18 months)

	Elasticity wrt u				HP-filtered Semi-elasticity wrt u			
Unemp. Duration	Occ Model	No Occ. Model I	No Occ. Model II	Data	Occ Model	No Occ. Model I	No Occ. Model II	Data
1-2 m	-0.432	-0.323	-0.260	-0.464	-0.168	-0.108	-0.093	-0.169
1-4 m	-0.314	-0.237	-0.183	-0.363	-0.178	-0.115	-0.101	-0.186
5-8 m	0.374	0.145	0.119	0.320	0.074	0.043	0.038	0.077
9-12 m	1.083	0.408	0.295	0.864	0.061	0.040	0.037	0.072
>13 m	1.787	0.734	0.484	1.375	0.044	0.031	0.026	0.044

functions are computed pooling the entire sample, while the distribution of incomplete spells between 1 and 18 months is calculated for each quarter and then averaged across quarters. Panel A of Table 25 shows that Model I generates about 50% more long-term unemployment (9-12 months) relative to the data. When considering durations between 13 and 18 months this discrepancy is even stronger, about 80%. At the same time, Panel B shows that the cyclical responses of the unemployment duration distribution generated by Model I are too small. It misses the semi-elasticity with respect to the unemployment rate by an average of about 40% across the entire distribution. This stands in contrast with the performance of the full occupational mobility model. Overall, Model I matches the unemployment survival functions by creating too dispersed unemployment durations within a typical period, in particular too many long spells, but the distribution then responds too little to the cycle.

Model II If one is willing to compromise on replicating the aggregate and age-group survival functions, however, the model without occupational mobility is able to generate larger cyclical volatilities. To show this we re-estimated the model by de-emphasising these survival functions. The esti-

Table 26: Targeted Moments. No Occupational Mobility II

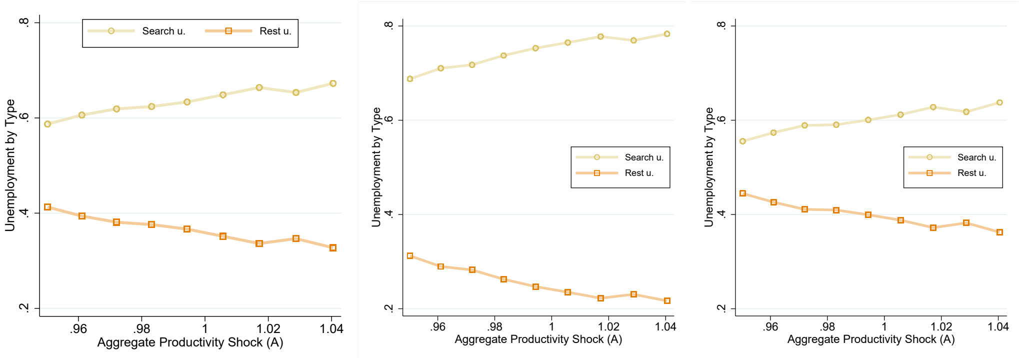
Moments	Model	Data	Moments	Model	Data	Moments	Model	Data
Aggregate Productivity			**U Survival w. Age**			**Returns to Human Capital**		
outpw	1.011	1.000	Young 2 months	0.710	0.697	5 years (OLS)	0.152	0.154
ρ_{outpw}	0.783	0.753	Young 4 months	0.421	0.381	10 years (OLS)	0.213	0.232
σ_{outpw}	0.0093	0.0094	Young 8 months	0.211	0.156			
Aggregate Matching Function			Young 12 months	0.128	0.073	**Empirical Separation moments**		
$\hat{\eta}$	0.341	0.500	Young 16 months	0.086	0.038	rel. sep rate young/prime	1.735	2.004
Unemployment Rate			Young 20 months	0.059	0.020	prob u within 3yrs for emp.	0.160	0.124
u	0.0335	0.0355				rel sep rate recent hire/all	4.167	4.945
U. Survival all workers			Prime 2 months	0.728	0.777			
agg. 2 months	0.724	0.758	Prime 4 months	0.445	0.485			
agg. 4 months	0.440	0.457	Prime 8 months	0.223	0.234			
agg. 8 months	0.220	0.208	Prime 12 months	0.132	0.137			
agg. 12 months	0.131	0.120	Prime 16 months	0.082	0.090			
agg. 16 months	0.082	0.076	Prime 20 months	0.053	0.061			
agg. 20 months	0.054	0.048						

mated parameter values are now $k = 102.019$, $b = 0.840$, $\eta = 0.245$, $\delta_L = 0.0047$, $\delta_H = 0.0016$, $\underline{z}_{corr} = 0.201$, $\rho_A = 0.997$, $\sigma_A = 0.0018$, $\rho_z = 0.993$, $\sigma_z = 0.0132$, $x_2 = 1.272$, $x_3 = 1.302$ and $\gamma_d = 0.0043$. In this case Model II exhibits a lower degree of search frictions as it estimates a lower value of k. Further, the overall dispersion of the z-productivity process is now much lower than in Model I and lower in the stationary distribution even than the occupational mobility models.

Table 26 shows that the unemployment survival functions of young and prime-aged workers are no longer well matched. In particular, Model II misses the distribution at longer durations for young workers and at shorter durations for prime-aged workers, such that age differences in job finding hazards have nearly disappeared. In contrast, we observe that the separation rate of young versus prime-aged workers is still significantly higher, though it remains below the targeted value. This version also displays a better persistence of workers' separation risk: recent hires have a 4 times higher separation rates than the average (vs. 5 in the data). The right panel of Table 22 shows that the main improvement of this version is that creates more cyclical volatility in the unemployment, job finding and separations rates as well as a stronger Beveridge curve.

Panel A of Table 25 - No Occ. Model II shows that although this model increases its cyclical performance, it still creates too much long-term unemployment (13-18 months) in the average quarter, where the proportion of long-term unemployed (among those with spells between 1-18 months) is missed by a large margin for all, young and prime-aged workers. Further, Panel B of Table 25 shows that this feature is also reflected in an unemployment duration distribution that does not respond as much to the unemployment rate as suggested by the data. Here we also find that Model II misses the semi-elasticity with respect to the unemployment rate by an average of about 40% across the duration distribution.[47]

[47] In these estimated one-sector versions of the model the behavior of spells with durations beyond 18 months might also impact the overall unemployment rate more than empirically warranted, especially when persistence creates a "first-in last-out pattern" when entering a recession with very long unemployment spells for those who lost their job early on in the recession. We focus on the distribution of spells up to 18 months because censoring issues in the SIPP restrict how accurately we can investigate the behavior of very long spells over the cycle.



(a) All workers (b) Young workers (c) Prime-aged workers

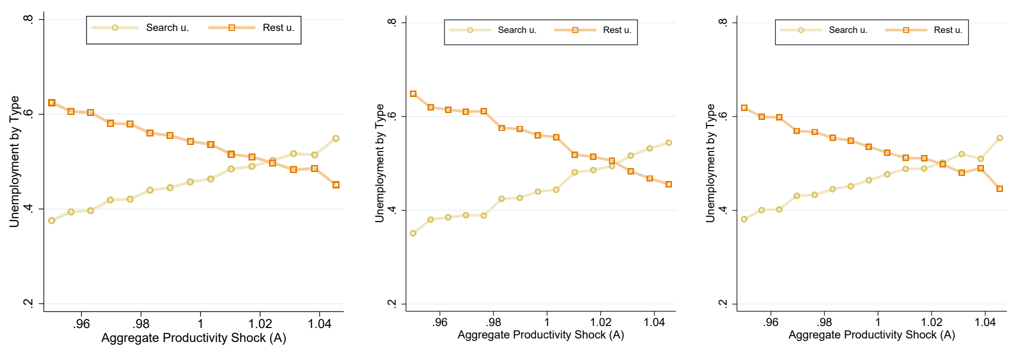
Figure 25: Unemployment decomposition - No occupational mobility, Model I

Discussion These calibrations show that without the z^r cutoff the no occupational mobility models cannot resolve the tension between individual unemployment outcomes and aggregate unemployment volatility. This arises as without the possibility of occupational mobility the new area of inaction is defined by the set of z-productivities that lie below the z^s cutoff down to the lowest value of the z-productivities. The cyclical response of this area now solely depends on $\partial z^s/\partial A$.

In the case of Model I, a less persistent and much more volatile z process creates enough heterogeneity in unemployment durations that allows it to match the empirical unemployment survival functions at the aggregate and across age groups. However, it also increases the heterogeneity in z-productivities relative to the cyclical range of A. This dampens the model's cyclical performance as it implies less responsive z^s cutoffs relative to the workers' z distribution, weakening the cyclical responses of job separations and the rate at which workers leave the area of inaction. Moreover, with a larger vacancy posting cost, Figure 25 shows that search unemployment is now more prominent than rest unemployment at any point of the cycle. Larger search frictions imply larger surpluses and therefore further reducing the cyclical responsiveness of the model.

The increased cyclical performance of Model II arises as its estimated z process becomes more persistent and less volatile, creating more responsive z^s cutoffs leading to stronger cyclical responses in job separations, as well as much more episodes of rest unemployment over all values of A. Figure 26 shows that rest unemployment episodes are now more prominent than search unemployment episodes even during economic recoveries. It is only for the highest values of A that search unemployment is more prominent, but only by a relatively small margin. With more responsive z^s cutoffs, an aggregate shock can now move somewhat larger masses of workers from rest into search unemployment episodes creating more amplification. This result is in line with Chassamboulli (2013), who extends the Mortensen and Pissarides (1994) model by adding permanent productivity differences among workers and shows that this feature allows that model to increase it cyclical performance relative to the data. However, Table 25 demonstrates that this comes at the cost of not being able to match the distribution of unemployment durations nor the dynamic behavior of this distribution over the cycle.[48] Thus, in Model II the average unemployment rate responds more to the cycle as aggregate

[48] An extrapolation from the above discussion appears to suggest a model with permanent heterogeneity, where all moves in and out of rest unemployment would be because of aggregate productivity changes, cannot resolve the tension



(a) All workers (b) Young workers (c) Prime-aged workers

Figure 26: Unemployment decomposition - No occupational mobility, Model II

productivity takes on a more prominent role in shaping the amount of rest unemployment, but the individual unemployment outcomes that underlie these dynamics become counterfactual.

The possibility of occupational mobility would have given workers in rest unemployment episodes another margin through which they can escape the area of inaction and get re-employed faster. As a result, the individual-level unemployment duration dependence (given an aggregate state) is not only affected by z^s but also by the distance to z^r. Further, since the z^r cutoff is at different distances from the z^s across the cycle, it creates a more cyclically responsive area of inaction. Overall, the addition of a more responsive area of inaction with respect to both worker heterogeneity and to the business cycle resolves the tensions documented above, even if we constrain the model to be consistent with the facts on occupational reallocation of the unemployed, as documented in Section 2 of the main text.

Supplementary Appendix

The Supplementary appendix can be found by following the coloured links to
https://www.dropbox.com/s/2crc29wqzqhbvr9/Supp_appendix_April_2020.pdf?dl=0.

References

[1] Abowd, J. and A. Zellner. 1985. "Estimating Gross Labor-Force Flows". *Journal of Business & Economic Statistics*, 3 (3): 254-283.

[2] Carrillo-Tudela, C. and L. Visschers. 2013. "Unemployment and Endogenous Reallocations Over the Business Cycle ". IZA Working Papers No. 7124.

[3] Campanelli, P., K. Thomson, N. Moon, and T. Staples. 1997. "The Quality of Occupational Coding in the United Kingdom". In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey Measurement and Process Quality*, 437-453. New York: Wiley.

[4] Chassamboulli, A. 2013. "Labor-Market Volatility in a Matching Model with Worker Heterogeneity and Endogenous Separations". *Labour Economics*, 24: 217-229.

between cyclical performance and fitting the unemployment duration distribution moments. In principle our estimation allows and has evaluated in the estimation procedure parameter tuples with a near-permanent z-productivity process (i.e. a persistence approximating 1). However, such parameter tuples yield stronger deviations from the individual-level unemployment outcomes we targeted compared to Model II.

[5] Hagedorn, M. and I. Manovskii. 2008. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited". *American Economic Review*, 98 (4): 1692-1706.

[6] Hausman, J., J. Abrevaya, and F.M. Scott-Morton. 1998. "Misclassification of the Dependent Variable in a Discrete-Response Setting", *Journal of Econometrics*, 87: 239-269.

[7] Hill, D. 1994. "The Relative Empirical Validity of Dependent and Independent Data Collection in a Panel Survey". *Journal of Official Statistics*, 10 (4): 359-380.

[8] Jäckle, A. and S. Eckman. 2019. "Is that still the same? Has that Changed? On the Accuracy of Measuring Change with Dependent Interviewing". *Journal of Survey Statistics and Methodology*, *forthcoming*.

[9] Jäckle, A. 2008. "Dependent Interviewing: Effects on Respondent Burden and Efficiency of Data Collection". *Journal of Official Statistics*, 24 (3): 1-21.

[10] Kambourov, G. and I. Manovskii. 2009. "Occupational Specificity of Human Capital". *International Economic Review*, 50 (1): 63-115.

[11] Kambourov, G. and I. Manovskii. 2008. "Rising Occupational and Industry Mobilty in the United States: 1968-97". *International Economic Review*, 49 (1): 41-79.

[12] Keane, M. and K. Wolpin, 2011. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment". *International Economic Review*, 42 (4): 1051-1103.

[13] Lynn, P. and E. Sala. 2007. "Measuring Change in Employment Characteristics: The Effect of Dependent Interviewing". *International Journal of Public Opinion Research*, 18 (4): 500-509.

[14] Mathiowetz, N. 1992. "Errors in Reports of Occupations". *Public Opinion Quarterly*, 56: 332-135.

[15] Magnac, T. and M. Visser. 1999. "Transition Models with Measurement Errors", *The Review of Economics and Statistics*, 81 (3): 466-474.

[16] Mortensen, D. and E. Nagypal. 2007. "More on Unemployment and Vacancy Fluctuations". *Review of Economic Dynamics*, 10: 327-347.

[17] Mortensen, D. and C. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment". *Review of Economic Studies*, 61(3): 397-415.

[18] Moscarini, G and K. Thomsson. 2007. "Occupational and Job Mobility in the US". *Scandinavian Journal of Economics*, 109 (4): 807-836.

[19] Poterba, J. and L. Summers. 1986. "Reporting Errors and Labor Market Dynamics". *Econometrica*, 54 (6): 1319-1338.

[20] Roys, N. and C. Taber. 2017. "Skill Prices, Occupations and Changes in the Wage Structure". Mimeo. Department of Economics, Royal Holloway, University of London.

[21] Shimer, R. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies". *American Economic Review*, 95 (1): 25-49.

[22] Sullivan, P. 2010. "Empirical evidence on occupation and industry specific human capital". *Labour Economics*, 17 (3): 567-580.

[23] Sullivan, P. 2009. "Estimation of an Occupational Choice Model when Occupations are Misclassified". *Journal of Human Resources*, 44 (2): 495-535.