



Working Papers

www.cesifo.org/wp

Distributional Implications of Joint Tax Evasion

Odd E. Nygård
Joel Slemrod
Thor O. Thoresen

CESIFO WORKING PAPER NO. 5915
CATEGORY 1: PUBLIC FINANCE
MAY 2016

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:* www.CESifo-group.org/wp

ISSN 2364-1428

Distributional Implications of Joint Tax Evasion

Abstract

Both buyers and sellers of goods and services may benefit from letting their economic transactions go unrecorded for tax purposes. The supplier reduces his tax burden by underreporting income, whereas the consumer gains from buying a non-taxed lower-priced product. The distributional implications of such joint tax evasion depend on the amounts evaded, on where the evaders on both sides of the market are found in the income distribution and how the financial gain is split between the suppliers and demanders. We use various data sources to identify tax evasion among sellers and buyers of goods and services. Results clearly suggest that the tax-evasion-controlled estimate of income inequality in Norway exhibits more income dispersion than official estimates.

JEL-Codes: D310, D630, H260.

Keywords: tax evasion, income inequality, expenditure approach.

Odd E. Nygård
Research Department
Statistics Norway
PO box 8131, Dep.
Norway – 0033 Oslo
oen@ssb.no

Joel Slemrod
Stephen M. Ross School of Business &
Department of Economics
University of Michigan
701 Tappan Street
USA – Ann Arbor, MI 48109
jslemrod@umich.edu

Thor O. Thoresen
Research Department
Statistics Norway
PO box 8131, Dep.
Norway – 0033 Oslo
tot@ssb.no

May 3, 2016

Financial support from the Norwegian Research Council is gratefully acknowledged. Elin Halvorsen and Bård Lian are acknowledged for helpful data advices and data assistance. We are grateful for comments by seminar participants of the Oslo Fiscal Studies Staff Seminar (Oslo, October, 2014), the Norwegian Research Forum on Taxation (Bergen, September, 2015), Nordic Workshop in Tax Policy and Public Economics (Oslo, November, 2015) and the 9th Norwegian-German Seminar on Public Economics (Munich, November, 2015).

1 Introduction

Who gains more from tax evasion, the rich or the poor? This is a complicated question because tax evasion generates numerous effects on factor and commodity prices (Kesselman, 1989; Slemrod, 2007; Alm and Finlay, 2013). Even when ignoring general equilibrium effects and discussing first-order distributional effects only, there are considerable complications to identifying effects, as we often lack reliable information on tax evasion. In the present paper we address how to account for the fact that much tax evasion behavior involves the participation of more than one taxpayer, and estimate the distributional consequences using data from Norway.

The point of departure here is that tax evasion often takes place in the interaction between buyers and sellers of services and commodities to households, referred to as collusive, or joint, tax evasion (Abraham et al., 2015). For example, the building and construction industry sector is one of the most tax evasion-ridden industries in Norway (KRISINO, 2011). Under a mutual agreement between sellers and buyers of goods and services, transactions or parts of transactions are often not reported to the tax authorities.

Distributional aspects of tax evasion have received little attention in the literature. Two exceptions are Bishop, Formby, and Lambert (2000) and Johns and Slemrod (2010), which both use micro data to address how measures of tax redistribution and income inequality are altered by accounting for tax evasion. A novelty of the present study is that we discuss distributional effects of tax evasion from a market perspective, where participants both of the supply side and the demand side may gain economically.

Measures of economic gains of tax evasion of the producers' side are obtained by using the so-called expenditure approach method (Pissarides and Weber, 1989), which belongs to an "indirect-method" way of identification of evasion behavior. See surveys of the literature in Andreoni, Erard, and Feinstein (1998), Schneider and Enste (2000), Slemrod and Yitzhaki (2002), Torgler (2007), Slemrod and Weber (2012), and Alm (2012). In an indirect-method approach, evasion is not measured directly, but indirectly via measurable traces of true income. Given that several groups of wage earners have limited scope for tax evasion (because third-party reporting of income is a standard procedure), and therefore represents a convincing benchmark, food consumption and income are compared between three groups: wage earners with little possibilities for tax evasion, wage earners who may be involved in tax evasion at the supply side (as employed craftsmen), and the self-employed. As in Pissarides and Weber (1989), we assume that there is a common slope in the Engel curves for food, but intercepts may differ in the three groups. In this way, by exploiting data from the Survey of Consumer

Expenditure (Holmøy and Lillegård, 2014), we can estimate the amount by which reported income must be scaled up in order to obtain true income levels for tax evaders,¹ and allow for variations in the degree of underreporting along the income range.

With respect to the buyers' side, we rely on a "direct method" for measuring tax evasion, by exploiting Norwegian survey data about purchases of illegal services (TNS Gallup, 2006; Opinion, 2009). We estimate a probability of being involved in transactions not reported to the tax authorities, dependent on characteristics, such as income and education. As we do not observe evaded amounts among buyers of the services, we benefit from our market transaction approach, and use non-reported income of the supply side to set the overall amounts.

The parameter estimates from the two econometric approaches are then combined to shed light on the distributional effects of joint tax evasion. Using the tax-benefit model LOTTE (Aasness, Dagsvik, and Thoresen, 2007) facilitates translating evasion behavior from both sides of the market into economic gains, and keeping track of the balance between amounts at the supply and demand side. By piecing together empirical evidence obtained from the supply side and the demand side of the market, we discuss two aspects of distributional effects of tax evasion: what the tax-evasion-controlled distribution of income looks like, and how the tax savings from not reporting income and transactions are distributed.

A measure of income inequality that controls for tax evasion accounts for the unreported income of suppliers and the benefits of consumers from paying a lower price. We obtain tax-evasion-adjusted incomes for suppliers by applying estimates of the expenditure approach directly. In the calculation of additions to income for consumers, we use the estimated evasion probabilities, but must also make assumptions about how prices in the hidden market deviate from prices in the regular market. The overall effect depends on how both the suppliers and consumers are positioned in the distribution of income. By this we are able to discuss how the "hidden-economy-controlled" income distribution compares to the official one: is it less or more equal?

The distribution of the reductions in tax burdens of suppliers and con-

¹The so-called "expenditure approach", set forth by Pissarides and Weber (1989) and exemplified by an application on British data, has sparked tax evasion examinations in several other countries, but estimates for Norway have so far not been provided. Tax evasion estimates for other countries by this method include Schuetze (2002) for Canada, Johansson (2005) for Finland, Engström and Holmlund (2009) and Engström and Hagen (2015) for Sweden, Martinez-Lopez (2013) for Spain, Paulus (2015a) for Estonia, Feldman and Slemrod (2007) and Hurst, Li, and Pugsley (2014) for the U.S, and Kim, Gibson, and Chung (2016) for Korea and Russia.

sumers due to tax evasion are, of course, related to these adjustments in income. However, the measurement of distributional effects becomes more complicated when we address reductions in tax burdens (tax savings), instead of income. Now, even more than for the income distribution perspective, results depend on how suppliers and consumers divide the economic gains, which we do not observe. For example, we may have a situation where the returns of the suppliers are squeezed to the extent that consumers retain all of the economic advantage, and the distribution of the economic gain is determined by the demand side alone.

Although the present study does not provide definitive evidence about all the elements involved in a robust depiction of the distributional effects of joint tax evasion, we show how the empirical evidence can be used to provide empirical illustrations of effects, given the conceptual foundations. The paper is organized as follows. Section 2 summarizes some of the main perspectives on tax evasion and collusive tax evasion in particular. In Section 3 we probe deeper into the theoretical background for our empirical investigations, whereas Section 4 presents the empirical approaches to obtain measures of economic gains for the supply and demand side, respectively, and estimation results. The overall effects on the distribution of economic well-being are summarized in Section 5, and Section 6 concludes the paper.

2 Preliminaries

From the perspective of the supply side (Allingham and Sandmo, 1972; Yitzhaki, 1974; Andreoni, Erard, and Feinstein, 1998), the agent has an (exogenously given) income level, and faces a tax rate. Then he decides how much income he will report to the tax authorities, comparing the expected utility of being detected and paying a penalty for tax evasion to the expected utility from being able to keep the evaded income. The same expected utility reasoning can be used to explain behavior at the demand side (Cremer and Gahvari, 1993), as exogenously given disposable income can be used to buy commodities or services when there are two types of possible transactions, regular and hidden.

Although we shall proceed from this standard framework, we acknowledge that these simple models do not provide a complete description of everyone's decision-making. One key criticism is that some taxpayers are likely not motivated by narrow self-interest alone but instead act as a member of a group, influenced by norms, customs, reciprocity, and patriotism. Further, individuals are obviously not only motivated by economic factors; notions such as shame, guilt and morality arguably also influence decisions in some circum-

stances. Others argue that the expected utility model does not provide a satisfactory description of peoples' perception of risk, i.e. they seem to overweigh low-probability events, which have resulted in contributions applying prospect theory (Dahmi and al-Nowaihi, 2007) and the rank-dependent expected utility model (Eide, Simson and Strøm, 2011). We believe, however, that the Allingham-Sandmo deterrence model explains the essential reasoning underlying the theoretical framework.

There are other studies challenging the predominant perspective in the literature that tax evasion is an interaction between a single economic agent and the government. For example, previous studies have elaborated on collusive tax cheating between employees and the employer, see Yaniv (1988; 1992) and Kleven, Kreiner, and Saez (2016). Boadway, Marceau, and Mongrain (2002) construct a model in which tax evasion requires the collaboration of at least two taxpayers. Using a game-theoretic approach, they describe how sanctions of tax evasion may lead to a direct increase in the expected cost of a transaction in the illegal sector, but may also increase the ability of an agent to commit to cooperate in tax evasion, and can therefore lead to more tax evasion. Similarly, Chang and Lai (2004) model collaborative tax evasion between a seller and his customer as a game, and incorporate a social norm into such collusive tax-evading activities. More prevalent tax evasion deteriorates social norms, penalties may induce more collaboration and may therefore increase tax evasion if tax evasion is already widespread, explained by a snowballing effect (or a critical-mass force). Abraham et al. (2015) show that, in a laboratory setting, the tax compliance norm has a stronger negative effect on the magnitude of collusive tax evasion than on independent tax evasion. Ognedal (2015), however, focuses on honesty being a competitive disadvantage and "tax morale" representing a bad substitute for sanctions in markets. Honesty reduces cheating, but the output may be less efficiently produced and less efficiently allocated between buyers.

The market-transaction perspective of the present study implies emphasis on the general equilibrium effects of tax evasion. Thus, the discussion of tax evasion not only accounts for effects working through different sides of the market, but in principle can control for a whole range of reactions by individuals and firms. Persson and Wissén (1984) study, analytically, the conditions under which the actual income distribution is more equal, or more unequal, than the distribution based on reported income. Richer descriptions of incidence effects of tax evasion can be obtained by employing computable general equilibrium (CGE) models, as in Alm and Sennoga (2010), who examine how much of the initial benefit of income tax evasion is retained by the evaders and how much is shifted via factor and commodity price changes

stemming from mobility.²

As the present study analyzes micro data for the supply and demand side, previous studies that use micro data to discuss distributional aspects of tax evasion are relevant. Bishop, Formby, and Lambert (2000) and Johns and Slemrod (2010) use data from the comprehensive random audit programs of the Internal Revenue Service (IRS) of the U.S., which allow the researchers to observe income as reported and as adjusted by an audit. Bishop et al. (2000) find that including unreported income has only a very small (negative) impact on pre-tax income inequality as measured either by the standard Gini coefficient or the extended Gini coefficient. Including both unreported income and additional taxes owed also has a negligible impact on inequality. Johns and Slemrod (2010) find that accounting for tax noncompliance makes the true income distribution more unequal, but the tax system becomes more progressive. This follows because a given percentage reduction in taxable income corresponds to a particularly high percentage reduction in tax liability for taxpayers with taxable income just above the taxpaying threshold.³ Kleven et al. (2011) also use variation in auditing to identify tax evasion magnitudes and, even though distributional effects are not a main topic of the paper, they report (p. 673) that those with relatively little self-reported income evade more, as a share of self-reported income, than those with relatively high self-reported income.

Further, Pashardes and Polycarpou (2008) employ an expenditure approach technique, outlined in Lyssiottou, Pashardes and Stengos (2004),⁴ and data from Cyprus to estimate tax evasion. Their findings suggest that the income underreporting biases estimates of both inequality and poverty downwards. Tedds (2010) uses an alternative way of implementing the expenditure-based method: parametric restrictions are relaxed and a nonparametric approach to the measurement of income underreporting is explored, thereby reducing the number of assumptions required for estimation. The approach is illustrated by estimating the effect of the Canadian Goods and Services Tax (GST) on income underreporting among self-employed, and the analysis concludes that the GST increased tax noncompliance by those with larger

²See also the discussion in Alm and Finlay (2013).

³Christian (1994) also analyzes data from the Taxpayer Compliance Measurement Program to discuss distributional aspects of tax evasion. The study finds that low-income individuals evade more than high-income individuals in the US. In 1988, taxpayers with (auditor-adjusted) incomes over \$100,000 on average reported 96.6 percent of their true incomes to the tax authorities, compared to just 85.9 percent for those with incomes under \$25,000.

⁴Instead of using expenditures on food only, as in Pissarides and Weber (1989), Lyssiottou et al. (2004) use information on a whole range of consumer goods.

amounts of self-employment income, whereas tax noncompliance by those with small amounts of self-employment income was not affected.

Finally, we note the results from studies discussing distributional effects of tax evasion by "discrepancy methods", meaning that data from an income survey are compared to the reported income of the income tax returns. Taxpayers may conceal part of their income from tax authorities, but might consider declaring a higher figure to an anonymous interviewer. Fiorio and D'Amuri (2005), Matsaganis and Flevotomou (2010) and Benedek and Lelkes (2011) use this method on data from Italy, Greece, and Hungary, respectively. Fiorio and D'Amuri (2005) find that the share of unreported income in Italy falls with income, Matsaganis and Flevotomou (2010) suggest tax evasion gives higher income inequality, more poverty and lower progressivity of the income tax, which is also in line with the findings in Benedek and Lelkes (2011).

3 Theoretical framework

Before we enter into the empirical investigation, we introduce some key conceptual terms by developing a simple theoretical framework. We assume there are two commodities in the economy: a numeraire good, c_1 , that cannot be sold in the informal market and a service, c_2 , that may be sold informally. An individual may be both a supplier and consumer of good c_2 , although we will refer to suppliers and consumers as if they are separate individuals.

3.1 Supply side

Each supplier has a skill level denoted n , and a skill type m , ($m \in [1, 2]$), φ_{c_1} and φ_{c_2} , such that they can supply c_1 or c_2 , but not both. Let s be equal to one if a person has skill type φ_{c_2} , and zero otherwise. A supplier of type φ_{c_2} decides whether to report the income for tax purposes, remit tax at rate τ , and thereby supply the service formally or not report the income and supply the service in the informal market.⁵ ⁶ In the regular market he gets the price before indirect taxation (before VAT), p_r , for the supply of c_2 , while the price is p_h in the informal market. Let x_r and x_h be the true before-tax income if operating in the regular or hidden economy, and let $x_r = np_r l_r$ and $x_h = np_h l_h$, where l_r and l_h denote the optimally chosen labor supply in each

⁵We assume for simplicity that both the seller and the buyer know with certainty whether a transaction will be reported for tax purposes.

⁶We can think of this as a decision in accordance with the Allingham and Sandmo (1972) framework.

sector. To see how this unreported income affects the income distribution, let us start by defining an elasticity, d_{ns} , characterizing how the probability of being of skill type φ_{c_2} changes when the skill level increases. We also define

$$k = \frac{y^*}{y}, \quad (1)$$

where y is the total after-tax reported income and y^* the total "true" after-tax income such that k measures the ratio of "true" income to reported income. Further, we define d_{ky} as the income elasticity characterizing the change in k when income y increases. These elasticities, like Feldstein's "distributional characteristics" (Feldstein, 1972), can be used to identify two effects on the income distribution from unreported income, depending on the sign and size of the two: i) if $d_{ns} > 0$, given the definition of s , the probability of being of type φ_{c_2} decreases with the skill level. Hence, we will find more of this type at the bottom of the income distribution and adjusting the income distribution for unreported income would tend to make it more equal, and ii) if $d_{yk} < 1$, underreported income as share of the total income decreases with reported income, which means that adjusting for unreported income makes the distribution more equal, *ceteris paribus*.

The supplier's expected economic gain from the tax evasion behavior is established by computing his expected financial gains from supplying l_h in the hidden market relative to supplying the same amount in the regular market. Let $x_r^* = np_r l_h$ be the (hypothetical) income if the hidden hours of work instead were supplied in the regular market. The expected gain for a supplier of participating in the hidden market, SG , can then be expressed as

$$\begin{aligned} SG &= (1 - \rho)(x_h - x_r^*(1 - \tau)) + \rho(x_h(1 - \tau) - x_h\theta - x_r^*(1 - \tau)) \\ &= [x_h - x_r^*(1 - \tau)] - \rho x_h(\tau + \theta), \end{aligned} \quad (2)$$

where θ is the penalty levied if caught evading (as a proportion to unreported hours), ρ is the probability of the fraud being discovered and τ is the income tax rate. The term in the square bracket captures the income difference between entering the hidden market and being paid x_h and letting the same hours of work be part of reported pre-tax income, x_r . Since there is a risk of being prosecuted and fined, the expected costs of being caught, $\rho x_h(\tau + \theta)$, enters too. If there is a negligible risk of being discovered, and if the price in the hidden market equals the pre-tax price in the regular market ($p_h = p_r$), such that $x_r^* = x_h$ for the same amount of labor, the gain from tax evasion corresponds to the (hypothetical) income tax ($x_r^*\tau$) on evaded income, i.e., the tax burden if the service was delivered in the regular instead of in the hidden market. As the shape of the income tax rate schedule affects decisions, for

example, making tax evasion disproportionately more attractive for taxpayers at the high end of the income distribution,⁷ tax evasion may make the tax system (effectively) less progressive.

It follows that the relationship between p_r and p_h determines the division of the financial gain between the supply side and the demand side. At the extreme, if $p_r(1 - \tau) > p_h$, working in the hidden market generates lower income than in the regular market (for the same working hours). Crucially, we do not observe prices in the two markets, and we therefore do not know how the gain is split between the two sides of the market. In Section 5 we discuss how we proceed to account for this in the distributional analysis.

Analogous to (2), we can define the supplier gain ratio as $f = \frac{SG}{y^*}$, and a corresponding elasticity d_{fy} , which shows how the ratio changes when true income changes. If $d_{fy} > 1$, then the ratio increases with income, and the rich gain relatively more than the poor from income tax evasion.

3.2 Demand side

Each consumer maximizes expected utility and determines the optimal demand for commodity c_1 and service c_2 . He can choose to purchase c_2 either in the informal or in the regular market, with the optimal amounts in the two markets denoted c_h and c_r , respectively. If he buys services in the hidden market and the fraud is uncovered by the tax authorities, he pays a penalty,⁸ given by a fraction of the tax evaded, denoted κ . The probability of being caught is denoted η , and π is the tax rate on the service, say the VAT. The expected financial gain for the consumer, CG , from purchasing c_h in the hidden market instead of in the regular market is given by,

$$\begin{aligned} CG &= (1 - \eta)(c_h p_r(1 + \pi) - c_h p_h) + \eta(c_h p_r(1 + \pi) - c_h p_h(1 + \pi) - c_h p_h \kappa \pi) \\ &= [c_h p_r(1 + \pi) - c_h p_h] - \eta c_h p_h \pi(1 + \kappa). \end{aligned} \quad (3)$$

The term in the square bracket represents the financial gain of the hidden purchase, whereas the last term describes the penalty. If the risk of being caught is negligible, the gain as share of the total true income, or the budget share of the gain, is given by

$$b = \frac{c_h(p_r(1 + \pi) - p_h)}{y^*}. \quad (4)$$

⁷This depends on the assumed form of the penalty function; see Yitzhaki (1974).

⁸A recent court case in Norway attests to this type of penalties on buyers. A buyer of hidden cleaning services was fined 20,000 Norwegian kroner, and sentenced to a 30-day (probated) prison term. Other customers of the same cleaner were fined, too.

Note that if $p_r = p_h$, then consumers simply remit the sales tax, $c_h p_r \pi$. However, if $p_r > p_h$ they benefit even more, and vice versa. The expected gain can then be written as

$$CG = \Phi b, \quad (5)$$

where $\Phi = \Pr(c_h > 0)$, the probability for the individual entering the hidden market.

Next, let us define an elasticity d_{by^*} , which describes how the budget share of the hidden market gain, b , changes when income changes. If $d_{by^*} > 1$, the share increases with the income. Similarly, let the elasticity $d_{y^*\Phi}$, define how the probability of entering the market depends on income. Then, the distributional effects depend on both d_{by^*} and $d_{y^*\Phi}$, how likely one is to enter the hidden market, and how much is spent in that market. In Section 4 we discuss how to obtain empirical counterparts of these components.

Thus, the economic gains at the demand side come in the form of different actual price vectors for different consumers, some paying prices below the observable regular price. The true real income can then be seen as $\frac{y^*}{P}$, where P is a price index.⁹ If $\frac{y^*}{P}$ increases with income, the (real) true income distribution is more unequal.

3.3 Equilibrium

To summarize the theoretical outline of tax evasion from a two-sided market perspective, the economy has three markets, one for the c_1 good and two, the regular and informal, for the c_2 service. Prices in the regular and hidden market for c_2 equilibrate demand and supply such that $F_h(p_r, p_h, \tau, \theta, \rho) = D_h(p_r, p_h, \pi, \kappa, \eta)$ and $F_r(p_r, p_h, \tau, \theta, \rho) = D_r(p_r, p_h, \pi, \kappa, \eta)$, where $F(\cdot)$ and $D(\cdot)$ are the aggregate supply and demand, respectively. The shape of these curves will determine how the prices in the hidden market relates to the regular market, and how the gains are shared between the supply and demand side.

⁹To see this, begin by defining real income for an individual as $\frac{y^*}{P}$, where $P = \frac{p_h c_h + p_r c_r + c_2}{p_r(1+\pi)c_h + p_r c_r + c_2}$, i.e. P is an individual-specific price index, dependent on how much the person buys in the hidden market. This means that the true income is adjusted upwards as long as $p_h < p_r(1 + \pi)$ (because $P < 1$). However, we have

$$P = \frac{p_h c_h + p_r c_r + c_2}{p_r(1 + \pi)c_h + p_r c_r + c_2} = \frac{y^*}{y^* + (p_r(1 + \pi) - p_h)c_h},$$

which implies that $P = \frac{1}{1+b} \Rightarrow 1/P = 1+b$. Then we can write $\frac{y^*}{P} = y^*(1+b) = y^* + CG$, when neglecting the risk of getting caught.

A key component of our approach to tax evasion is that there is an equilibrium condition determining tax evasion on both sides of the hidden market. If we let the observations of supply and demand side be represented by i and j , respectively, this can be seen as

$$\sum_i x_{hi} = \sum_j p_h c_{hj}, \quad (6)$$

i.e., the sum of evaded income equals the sum of hidden consumer payments.¹⁰ However, as we observe x_{hi} , and only establish to what extent $c_{hj} > 0$ with respect to the demand side, we use Equation (6) to obtain aggregate evaded amounts for the demand side, too.

In the next section we turn to the estimation of key parameters of this framework, which are used to describe the overall distributional effects of tax evasion. We estimate the parameter k of the supply side, and to some extent let it vary according to income. By establishing how k relates to the income level, and placing the evaders of the supply side in the income distribution, we address the size of d_{ky} and d_{sy} . Similarly, we provide evidence on how expenditures on hidden goods relate to income, thus discussing $d_{\Phi y^*}$ and d_{by^*} .

4 Quantifying tax evasion in both sides of the market

4.1 Identification of the evaded income of suppliers

The expenditure approach follows from the assumptions that some individuals have the opportunity to underreport, referred to as skill type φ_{c_2} in Section 3.1, while others do not, and that the groups have similar preferences for a consumption good. It is assumed that for everyone consumption is determined by true permanent income, y^* , and a number of individual control variables, Z' . When using the log form, we have the following Engel curve relationship, $\ln e = Z\gamma + \beta \ln y^*$, where γ and β are parameters.¹¹ Pissarides and Weber (1989) let e be represented by expenditures on food and assume that the self-employed are the only group that has scope for underreporting.

As already seen in Equation (1), the relationship between observed income, y_{im} , and true, permanent income, y_{im}^* , for individual i of type m can

¹⁰We suppose that they are subject to regular prices and taxes on factor of inputs.

¹¹Thus, we assume a log-linear Engel curve. One alternative is to employ a quadratic form, as argued for by Banks, Blundell, and Lewbel (1997), Lyssiotou, Pashardes, and Stengos (2004), and Fortin, Lacroix, and Pinard (2010).

be described by a proportionality factor, k_{im} , where we assign type m to indicate that there are differences across individuals in the scope for evasion, referred to as type 1 and 2 in the previous section. Thus, $k_{im} (> 1)$, shows the factor by which the observed income for individual i of type m must be multiplied in order to obtain true income. Influenced by Pissarides and Weber (1989) and several subsequent papers adopting their line of research, we let one of the groups (m) be represented by the self-employed. However, we shall also allow that some groups of wage earners are involved in tax evasion too, thus acknowledging that some wage-earning groups, such as painters or carpenters, may use their "leisure time" to work in the informal economy.¹² See also Dunbar and Fu (2015) and Paulus (2015a).

Further, following Pissarides and Weber, standard applications of the expenditure approach assume that observable income fluctuates around permanent income by a factor g in the two groups, seen as $y_{im} = gy_{im}^*$, and usually assume that the coefficients $\ln g_{im}$ and $\ln k_{im}$ are lognormally distributed around their means, $\ln g_{im} = \mu_{gm} + u_{im}$ and $\ln k_{im} = \mu_{km} + v_{im}$. Then it follows that the relationship between (true) permanent income and observable income is $\ln y_{im}^* = y_{im} - (\mu_{gm} - \mu_{km}) - (u_{im} - v_{im})$, generating the following Engel curve,

$$\ln e_{im} = Z_{im}\gamma + \beta \ln y_{im} - \beta (\mu_{gm} - \mu_{km}) + \beta (u_{im} - v_{im}) + \varepsilon_{im}. \quad (7)$$

If for expositional reasons we let the self-employed, in contrast to the wage earners, be defined as evaders, the Engel curve is adjusted by an indicator variable, q_i , which takes the value 1 for the self-employed, SE . When letting the (non-evading) salary workers be symbolized by SW , Equation (7) becomes

$$\begin{aligned} \ln e_{im} = & Z_{im}\gamma + \beta \ln y_{im} + \beta (\mu_{kSW} - \mu_{gSW}) \\ & + \beta q_i [(\mu_{kSE} - \mu_{kSW}) - (\mu_{gSW} - \mu_{gSE})] \\ & + \beta (u_{im} - v_{im}) + \varepsilon_{im}. \end{aligned} \quad (8)$$

The mean of k_{SE} is given by $\ln \bar{k}_{SE} = \mu_{kSE} + \frac{1}{2}\sigma_{vSE}^2$, where σ_{vSE}^2 is the variance of v_{im} for $m \in SE$. Also, as $k_{SW} = 1$ for $m \in SW$, and as the income means in the two groups are identical, $\ln \bar{g}_{SE} = \ln \bar{g}_{SW}$, Equation (8) can be rearranged into the following reduced form, which is the standard empirical specification used to obtain estimates of k ,

$$\ln e_i = Z_i\gamma + \beta \ln y_i + \delta q_i + \xi_i. \quad (9)$$

¹²For example, they may provide paid help to family or acquaintances. As emphasized by Williams (2008), many informal economy buyer-seller interactions are of this type.

As $\delta = \beta [\mu_{kSE} + \frac{1}{2}(\sigma_{uSE}^2 - \sigma_{uSW}^2)]$ and $\xi_i = \beta(u_{im} - v_{im}) + \varepsilon_{im}$, an estimate of the adjustment factor k is given by

$$\bar{k}_{SE} = \exp \left[\mu_{kSE} + \frac{1}{2}\sigma_{vSE}^2 \right] = \exp \left[\frac{\delta}{\beta} + \frac{1}{2}(\sigma_{vSE}^2 + \sigma_{uSW}^2 - \sigma_{uSE}^2) \right]. \quad (10)$$

However, as σ_{vSE}^2 , σ_{uSW}^2 , and σ_{uSE}^2 are usually not known, a standard empirical approach, along the lines of Pissarides and Weber, involves obtaining estimates of the variance of the residuals, ζ_{im} , from an expression, $\ln y_{im} = B_{im}\psi + \zeta_{im}$, where B_{im} includes a set of instruments for permanent income. Thus, an estimate of \bar{k}_{SE} is obtained by¹³

$$\bar{k}_{SE} = \exp \left[\frac{\delta}{\beta} \pm \frac{1}{2}(\sigma_{\zeta SE}^2 - \sigma_{\zeta SW}^2) \right]. \quad (11)$$

Our approach to obtaining estimates of k for specific groups of salary workers follows the same type of reasoning. In our main specification we estimate Equation (9) directly by using a measure of permanent income for y_i , which simplifies Equation (11). We shall return to this shortly.

Moreover, as a key objective here is to obtain information about how k varies with respect to income, corresponding to obtaining information about d_{yk} (see Section 3.1), we introduce non-linearities in the measurement of δ . This can straightforwardly be done by introducing a dummy variable denoting high income, hi_i , and letting it be interacted with the variable exhibiting that there is scope for underreporting, q_i , in Equation (9),

$$\ln e_i = Z_i\gamma + \beta \ln y_i + \delta q_i + \lambda(q_i \times hi_i) + \nu_i, \quad (12)$$

where ν_i is the error term.

4.2 Data and estimation results for expenditure approach

Estimates of k are obtained by examining data from the Norwegian Survey of Consumer Expenditure (Holmøy and Lillegård, 2014). We pool information over the time period from 2003 to 2009, so that we have consumer expenditure information for more than 4,000 observations. The estimation is carried out at the household level, as this is the relevant economic unit for consumer expenditures. This raises questions about how individual information is represented at the household level. All these questions, including the procedure to assign permanent income to individuals, are addressed in Appendix A.

¹³Pissarides and Weber discuss the result for both a lower bound case ($\sigma_{vSE}^2 = 0$), and an upper bound alternative ($\sigma_{uSW}^2 = \sigma_{uSE}^2$). See also Wangen (2005).

Measures of permanent income are linked to the expenditure data by using an income panel dataset for the whole population (Statistics Norway, 2005). As personal ID numbers are unavailable, we merge by using "backward identification" methods, exploiting that there is overlapping information in the two data sets. As we do not find a unique match for all observations, some observations are lost, but we are able to retain more than 70 percent of the sample. The measure of permanent income is obtained by taking averages of income over seven years.

Estimation results for eight different specifications are presented in Table 1. In the upper panel we present the standard expenditure approach results for the self-employed and for those wage earners assumed to have some scope for evasion, based on the specification in Equation (9). Results for specifications letting income be represented by both yearly income and permanent income are provided, using both ordinary least squares (OLS) and instrumental variables (IV) techniques in the estimations.¹⁴ Given that the seven-year income average represents permanent income adequately, Equation (11) is simplified, as $\sigma_{uSW}^2 = \sigma_{uSE}^2$, and the remaining contribution comes from variance in the self-employment underreporting rate (σ_{vSE}^2), and $\bar{k}_{SE} = \exp \left[\frac{\delta}{\beta} + \frac{1}{2} (\sigma_{\zeta SE}^2 - \sigma_{\zeta SW}^2) \right]$ is used to obtain estimates of \bar{k}_{SE} . The lower panel of the table presents results when allowing for income level interaction in the group shift variable (self-employed and specific groups of wage earners), as given by Equation (12).

Estimates of k show little variation depending on the choice of income measure and the estimation procedure. Estimates for the self-employed range from 1.20 to 1.25 and, when using the preferable permanent income specification, are 1.23 and 1.21, for OLS and IV estimations, respectively.¹⁵ Thus, the OLS and IV estimates are close.¹⁶ To our knowledge these estimates are the first self-employment tax evasion estimates for Norway using the expenditure approach; we find it reassuring that they are not far from estimates of the share of underreported income among the self-employed in Denmark

¹⁴As in most of the previous literature, including Pissarides and Weber (1989), estimates are also obtained by using IV methods. 2SLS estimations results are based on using the size of one's house, a dummy variable for higher education and (log of) capital income as instruments. Engström and Hagen (2015) report results that support using capital income as an instrument.

¹⁵In a companion paper we use charitable contributions instead of food consumption for identification, and find k -values very close to the estimates reported here (for a much larger sample).

¹⁶Note that the F-statistic of the first stage, p-values of the Sargan's overidentification test, and the Wu-Hausman endogeneity test all suggest good performance of the IV approach.

and Sweden, as reported by Kleven et al. (2011) and Engström and Hagen (2015), respectively.¹⁷

We see no evidence of tax evasion among wage earners on average. However, and crucially, given the perspective of the present study, when allowing for different effects by income range, we see indications of tax evasion among salary workers with reported income below the median, of between 8 and 10 percent. However, note that this estimate is not significant statistically different from 1. Similarly, when we differentiate between self-employed with high and low income, results suggest that there may be more tax evasion at low (reported) income levels; however, estimates of the interacted variable, λ_{SE} , are not statistically significant.¹⁸ These results do not accord with findings reported in Johns and Slemrod (2010) for the U.S., as they find that the ratio of aggregated misreported income to true income generally increases with income.

4.3 Buyers in the hidden market

As spelled out in Section 3.2, for evidence of tax evasion at the demand side we exploit information from two sample surveys that were carried out in 2006 and 2009 to enhance knowledge about the informal economy through interviews. Results are presented in TNS Gallup (2006) and in Opinion (2009). Of course, given the topic of the surveys, there is a clear potential for non-response bias. The response ratio is 0.56 in the 2009 survey and it appears to be somewhat lower in 2006 (although not reported). Correspondingly, the 2009 survey arguably replicates population distributions better than the 2006 survey.¹⁹ Here we pool the information from the two surveys, which implies that we exploit information from approximately 3,000 respondents, about 1,000 from the 2006 survey and 2,000 from the 2009 survey. More details about the data can be found in Appendix A.

In the surveys, 13 percent and 23 percent of the respondents, respectively for 2006 and 2009, report that they have bought services and/or goods in the hidden market during the previous two years. They are also asked

¹⁷Both Kleven et al. (2011) and Engström and Hagen (2015) report results in terms of the fraction of income, which means that recalculation is needed in order to compare with our estimates for k .

¹⁸Estimation of specification allowing for further disaggregation (quartiles) does not reveal any indications of distinctive differences between agents belonging to the two quartiles with income below the median. However, (again) no significant estimates are obtained.

¹⁹Obviously, there is a danger that the nonresponse contaminates results. However, note that there is little empirical support for the notion that low response rate de facto produce estimates with high nonresponse bias (Groves, 2006).

Table 1: Estimation results for expenditure approach. Pooled consumer expenditure data 2003–2009

	OLS, yearly	IV, yearly	OLS, permanent	IV, permanent
Slope, β	.322 (13.85)***	.332 (7.79)***	.401 (15.96)***	.374 (8.19)***
Self-employed, δ_{SE}	.070 (2.27)**	.060 (1.92)*	.067 (2.16)**	.056 (1.77)*
Salary worker, δ_{SW}	-.015 (-0.67)	-0.018 (-0.80)	-.004 (-0.18)	-.010 (-0.45)
Implied k_{SE}	1.25 (4.20)**	1.20 (3.06)*	1.23 (5.89)**	1.21 (4.14)**
R^2	.445	.446	.450	.452
1 st stage F-statistic		263.8		337.2
Sargan (p - value)		.230		.632
Wu-H (p - value)		.865		.344
With interaction, high vs. low income				
Slope, β	.325 (13.64)***	.335 (7.66)***	.409 (15.84)***	.380 (8.07)***
Self-employed, δ_{SE}	.065 (1.50)	.057 (1.27)	.082 (1.87)*	.062 (1.39)
High-inc, SE, λ_{SE}	.009 (0.14)	.006 (0.09)	-.029 (-0.47)	-.013 (-0.21)
Salary worker, δ_{SW}	.034 (1.01)	.029 (0.84)	.058 (1.72)*	.045 (1.31)
High-inc, SW, λ_{SW}	-.085 (-2.07)**	-.082 (-1.95)*	-.109 (-2.64)***	-.097 (-2.31)**
Implied $k_{SE,LI}$	1.22 (1.93)	1.19 (1.50)	1.27 (4.17)**	1.23 (2.67)
Implied $k_{SE,HI}$	1.26 (2.28)	1.21 (1.50)	1.19 (2.13)	1.19 (1.59)
Implied $k_{SW,LI}$	1.10 (0.78)	1.08 (0.54)	1.12 (1.67)	1.09 (0.90)
R^2	.445	.446	.451	.453
1 st stage F-statistic		253.9		324.9
Sargan (p - value)		.279		.595
Wu-H (p - value)		.856		.337
No of observations	4, 207	4, 055	4, 213	4, 057
$p < 0.1^*$ $p < 0.05^{**}$ $p < 0.01^{***}$				

Notes: t-statistics and chi-square statistics in parentheses for regression coefficients and k values, respectively. Chi-square statistics for k are based on the delta method, accounting for variance in β and δ . Regressions include controls for age, age squared, and number of adults and children. Add. instrum. IV regressions: size of house, dummy for higher education, and log of capital income.

about the amount spent on hidden services, although with respect to relatively wide intervals of expenditures. In Table A2 in the Appendix, we have pooled the information from both surveys and show how expenditures correlate with household income. The table shows that a clear majority declare that they have spent less than 50,000 Norwegian kroner (NOK) (over the last two years),²⁰ and the overall picture is that expenditures increase in income. However, we find the information on the amounts spent in the hidden market to be too inexact to be used in the present context. Thus, as already described by Equation (6), instead we use the relationship to the supply side to establish demand side evaded amounts. This means we postulate an assumption regarding an Engel curve for hidden expenditures (which will be varied to test for robustness) and use hidden market participation and other characteristics (including income) to place the evaders in the income distribution. By this empirical strategy we also enforce that the evaded amounts of the supply side (obtained by the expenditure approach) equal the hidden payments of the demand side. In the next section we shall return to the practical implementation of this condition, and what it means in terms of distributional effects.

To obtain information about the determinants for being involved in non-recorded purchases, we estimate a probit model,

$$\Pr(c_{hj} = 1 \mid Z_j), \quad (13)$$

where $c_{hj} = 1$ for individual j if he/she reports having paid for services in the hidden market, i.e., have $c_{hj} > 0$, where Z_j symbolizes control variables. By pooling observations of the surveys of 2006 and 2009, Equation (13) is estimated on a sample of approximately 3,250 observations; see descriptive statistics in Table A1 in Appendix A.

The estimation results reported in Table 2 show that income, as well as gender and region, are significant explanatory characteristics for the probability of participating in the hidden market. The probability of being in the informal market increases with income (although at a decreasing rate as indicated by the square term²¹), males have a higher probability than females, and location matters, as people in eastern part of Norway (including in the capital Oslo) have higher probability for being involved in hidden transactions. The positive relationship with income suggests that informal markets may increase inequality.

²⁰According to exchange rates for 2009: 1€ = 8.73 Norwegian kroner (NOK), and 1\$ = 6.64 NOK.

²¹The estimated effect of income on the probability of participating in the informal economy remains positive until income reaches 3.7 million Norwegian kroner.

Table 2: Probit estimation for buying services in the hidden market, pooled data, 2006 and 2009

	Coefficient
Gross household income	.0008 (4.54)***
Income squared	-2.14x10 ⁻⁷ (-3.30)***
Age	-.001 (-0.74)
Male	.135 (2.43)**
Education	.044 (0.73)
Self-employed	.045 (0.37)
Eastern region dummy	.287 (4.19)***
Constant	-1.524 (11.65)***
Likelihood Ratio	59.18
No. of observations	3,259

$p < 0.1^*$ $p < 0.05^{**}$ $p < 0.01^{***}$

Notes: Income measured in 1,000 Norwegian kroner. z-statistics based on standard deviation reported in parentheses

Table 3: Probability of buying services in the hidden market for three gross household income levels

200,000 NOK		800,000 NOK		1,400,000 NOK	
Male	Female	Male	Female	Male	Female
.105	.083	.185	.151	.243	.203

Note: Income measured in 2009-values

To depict how these estimates are used to compute the tax-evasion-corrected incomes, in Table 3 we show probabilities for three different household income levels, allowing for differences across gender.²² For example, we see that the probability of entering the informal market increases from 0.11 to 0.24 when the income of the household goes up from 0.2 million NOK to 1.4 million NOK.

Given that the supply side estimates have been obtained for households, whereas the estimates of Table 2 have been derived at the individual level, we let the individual actions of the demand side represent a "household prob-

²²Non-significant explanatory variables are not used in this calculation, and the region parameter is set to its average.

ability", in order to coordinate the two sources of information.

5 Distributional effects of tax evasion

5.1 Measurement of the distribution of tax evasion when information is limited

Recall that one of our central objectives is to obtain an estimate of the "hidden-economy-controlled" income distribution, and see how it relates to the official one. We obtain an income distribution adjusted for sellers' unreported income by, for each evading household, letting their disposable income be adjusted by the relevant k . As discussed in Section 3.2, from the sellers' side, a description of the true distribution of income should also reflect that some households pay lower prices on part of their consumption.

But, as discussed in the Introduction, and as made clear in the outline of the decision-making in Section 3, it is possible to adopt another angle in the discussion of distributional effects, namely by addressing information about the individual tax savings of not reporting income and transactions. This perspective brings the tax incidence challenge of the present analysis to the surface. Recall that we do not observe p_r and p_h , which implies that we do not observe how the gains are divided between sellers and buyers. We shall, as a point of departure and to fix ideas, assume that the consumer price in the informal market is equal to the pre-VAT price in the regular market, $p_r = p_h$, and the individual gain is defined by the indirect taxation, as the VAT. But in reality, for example, we may have the case where the returns of suppliers are squeezed to the extent that the purchasers receive all or most of the economic advantage, the saved income tax included. Of course, we can have the opposite situation too. We will return to assumptions regarding the splitting of the economic advantage between the two sides of the market in the sensitivity analysis. More generally, this raises the questions of who in reality bears the burden of the tax and what the real incidence of taxes looks like, but in depth discussions of this are beyond the scope of the present analysis.

Equation (2) characterizes the individual benefit of the evading supplier. Given that $p_r = p_h$, it follows that $\Delta_t = t^* - t$ can be used as a starting point, where t^* is the income tax burden corresponding to reporting all income truthfully, where the latter is defined by $x_r^* = x_r + \Delta_y$, where $\Delta_y = y^* - y$. Then, Equation (3) defines the advantage of buyers, and when $p_r = p_h$, and given the present VAT rate in Norway, which is 25%, it implies that $\frac{p_r(1+\pi)-p_h}{p_r(1+\pi)} = 0.2$; the price in the hidden market is 20 percent below the price

in the regular market.²³

As shown in Section 2, earlier contributions in the literature attest to the possibilities of applying different perspectives on the distributional impact of tax evasion. For example, the results reported in Johns and Slemrod (2010) illustrate that one may get results that show that tax evasion is increasing in income, but which still would give a more progressive tax schedule. In the following we shall describe the distributional effects from both angles; both as an adjustment to income, given that we are interested in how the distribution of "real" income looks like, and from a perspective where tax evasion gives reduced tax burdens.

We do this by piecing together the econometric evidence presented in Section 4. In order to translate the empirical findings from the previous section into distributional effects, we utilize the tax-benefit model LOTTE (Aasness, Dagsvik, and Thoresen, 2007). This allows us to impute our estimates into a representative sample of the population, to account for evasion behavior of both sides of the market, and to keep track of the balance between amounts at the supply and demand side, as seen in Equation (6).

When imputing the supply side unreported income, we use the OLS estimates for permanent income from column four of Table 1. Although estimates of k are not statistically significant different from each other, we let k differ with respect to point estimates. This means that income for households containing at least one self-employed is adjusted according to two estimates of k : 1.27 for below-median-income households, and 1.19 for those with observed income above the median. Correspondingly, an income correction has been allocated to low-income salary workers, conditional on the type of industry, based on the estimate of $k_{SW,LI}$ in Table 1. As we shall see, the main results are not contingent on this differentiation of k .

With respect to individuals as consumers, an important component of the empirical framework is to establish an Engel curve for the hidden service. When expenditure is given by $e_h = p_h c_h$, we assume that the hidden economy Engel curve is

$$\ln e_h = (a + \varphi \ln x), \quad (14)$$

where a is a constant, x is gross household income (see, for example, Table 2), and φ can be interpreted as the income elasticity, corresponding to d_{by^*} (see Section 3.2). Then, when using the estimation results of Section 4.3, the expected expenditure on a hidden market service for the household

²³In a price-negotiation situation, the customer may be offered the choice to pay with or without VAT, which then can justify equality of pre-tax prices, at least as a point of departure.

represented by the buyer j can be written as,

$$\Pr(c_{hj} = 1 \mid Z_j) \exp(a + \varphi \ln x_{hj}), \quad (15)$$

which if combined with the equilibrium condition in (6) gives

$$\sum_i x_{hi} = \sum_j \Pr(c_{hj} = 1 \mid Z_j) \exp(a + \varphi \ln x_{hj}), \quad (16)$$

which in turn can be used to get an expression for a ,

$$a = \ln \left(\frac{\sum_i x_{hi}}{\sum_j \Pr(c_{hj} = 1 \mid Z) x_{hj}^\varphi} \right), \quad (17)$$

for any given value of φ . As a point of departure, we set φ at 1, which means that budget shares of informal consumption, conditional on some positive consumption, are constant along the income scale. This is in accordance with findings described in a report from a Danish expert group (De Økonomiske Råd, 2011). Recall from Equation (5) that the expected gain consists of a budget share and the probability to purchase. Thus, we assume that there exists a hidden economy Engel curve, characterized by φ , where our empirical contribution comes from identifying the participants and their characteristics.

5.2 Effects on income inequality and distribution of tax burdens

We begin, in Table 4, by showing how disposable income should be adjusted due to the underreporting of sellers' income.²⁴ The table shows that true post-tax income, on average, is 12,772 NOK higher than reported income. In total, this corresponds to post-tax income being 29 billion NOK higher if income is correctly reported. The overall picture is that the rate of underreporting increases with income. In particular, the underreporting rates are relatively high in deciles 9 and 10 (see the last column of Table 4). Even though we found more tax evasion (as measured by k) among both self-employed and wage earners in the low end of their respective income distributions,²⁵ we find the opposite effects when the effects are calculated with respect to the overall income distribution. A key factor behind this result

²⁴We rank the households according to their "true" income. If we used the reported income concept, the tax evaders would be placed lower in the distribution than they actually are, as discussed in Johns and Slemrod (2010).

²⁵Recall that we used point estimates even though differences were not statistically significant. This means that an alternative empirical strategy, consisting of using only one estimate for k , would only strengthen results.

Table 4: Distribution of tax-evasion-adjusted income of suppliers, household post-tax income, 2009

Decile	Share of self-employed (%)	Reported post-tax income	Tax evasion adjusted post-tax income	Reported as share of adjusted post-tax income (%)
1	3.12	93,112	93,688	99.39
2	1.94	172,946	173,680	99.58
3	3.13	223,609	225,153	99.31
4	3.65	283,025	285,232	99.23
5	4.79	343,196	346,847	98.95
6	10.51	415,406	424,715	97.81
7	11.29	505,754	514,987	98.21
8	12.53	598,407	610,730	97.98
9	18.44	710,148	731,984	97.02
10	34.15	1,069,290	1,135,595	94.16
All	10.35	441,489	454,261	97.19
Gini coeff.		0.356	0.364	

Notes: Households ranked by tax-evasion-adjusted post-tax income in 2009.
All values refer to decile mean.

is the placement of the self-employed in the overall income distribution, as reported in the first column of the table, which shows that there are groups other than the self-employed that dominate the low end of the income distribution. It follows from this that the Gini coefficient for true income is higher than for reported income, as seen in the bottom last row of Table 4.

Next, in Table 5, we turn the attention to the reduced tax burdens among sellers and buyers following from participating in the hidden market. Thus, with respect to the suppliers, non-reported income that forms the basis for the income distributions reported in Table 4 is taxed, and the distribution of the tax burden difference between non-reported income being taxed or not (SG) is shown in the first column of Table 5. With respect to the consumers, as discussed in Section 5.1, the reduced tax burden (the consumer gain, CG) is equal to the indirect taxation of the hidden consumption (say the VAT), given that assumption that pre-tax prices of the hidden and the regular market are equal, $p_r = p_h$, and the reduced tax burden by buying in the hidden market is equal to the difference in after-tax prices ($p_r(1 + \pi)$ vs p_h). As explained in Section 3.2, our approach implies that the gains in terms of reduced prices are exactly equal to the reduced tax burden for the consumers.

Table 5 shows that the average size of CG is smaller than the average of SG distribution of CG . As expected, given the results of Table 2 and Table

Table 5: Distributions of reduced tax burden for suppliers and consumers participating in the hidden market

Decile	Supplier gain (<i>SG</i>)	Supplier gain, income share (%)	Consumer gain (<i>CG</i>)	Consumer gain, income share (%)
1	217	0.23	273	0.24
2	437	0.25	509	0.25
3	1,124	0.50	740	0.30
4	2,019	0.71	1,099	0.37
5	2,966	0.85	1,573	0.45
6	3,676	0.86	2,296	0.57
7	3,615	0.70	3,229	0.69
8	4,864	0.79	4,259	0.79
9	8,850	1.20	5,636	0.88
10	27,904	2.43	12,315	0.93
All	5,567	1.22	3,193	0.70
Gini coeff.	0.920		0.545	

Notes: Households ranked by tax-evasion-adjusted post-tax income in 2009. All values refer to decile mean. Engel elasticity and price assumption: $\varphi = 1$, $p_r = p_h$.
Income shares as share of tax-evasion-adjusted post-tax income.

3, tax evasion at the demand side also generates a more unequal "true" income distribution; see the last column of the table, where consumer gains are described in terms of shares of income. Table 5 shows that the reductions in tax burdens are rather modest both in an absolute and relative sense in deciles 1–3, among both suppliers and consumers. At the top of the income distribution, in contrast, the gains for both sides are large.

The joint supply and demand side effect of tax evasion on the income distribution is summarized in Table 6, where the distribution of tax-evasion-adjusted income, accounting for evasion at both sides of the market, is shown, and compared to an income distribution based on a conventional definition. Thus, the post-tax income of Table 4 has been added to the price gain of consumers in Table 5,²⁶ and turned into an overall tax-evasion-corrected measure of disposable income. We see that the Gini coefficient for this income concept is 0.365, which is notably higher than the Gini coefficient income inequality for conventional income of 0.356. Thus, our method clearly suggests that the "hidden-economy-controlled" income inequality is higher than the official one. Note also that the income inequality, as measured by the Gini coefficient, after the correction for supply-side tax evasion is 0.364, which

²⁶Which is identical to the reduced tax burden following from not paying the VAT.

Table 6: Conventional income distribution compared to income distribution controlled for tax evasion of suppliers and consumers, household post-tax income, 2009

Decile	Reported post-tax income	Post-tax income adjusted by collusive tax evasion	Reported as share of adjusted post-tax income (%)
1	93,112	93,911	99.15
2	172,946	174,121	99.33
3	223,609	225,830	99.02
4	283,025	286,290	98.86
5	343,196	348,427	98.50
6	415,406	427,131	97.25
7	505,754	518,539	97.53
8	598,407	615,565	97.21
9	710,148	738,510	96.16
10	1,069,290	1,146,215	93.29
All	441,489	457,454	96.51
Gini	0.356	0.365	

Notes: Households ranked by tax-evasion-adjusted post-tax income in 2009. All values refer to decile mean. Engel elasticity and price assumption: $\varphi = 1$, $p_r = p_h$.

suggests that the (additional) effect of the demand side is relatively modest, as measured here.

Finally, in Table 7 we show how the distribution of the tax revenue is affected by tax evasion, accounting for effects through both direct and indirect taxation. As shown, the compliance rate decreases with income, implying that the tax schedule in reality is less progressive than shown by official figures. The mean compliance rate for the whole population is 96.5 percent, which corresponds to a revenue loss of approximately 20 billion NOK, in total. The VAT loss is calculated to about 7 billion NOK, whereas the tax revenue from the personal income tax would have been approximately 13 billion NOK higher if the self-employed and (and wage-earner craftsmen) reported all their (true) income.

5.3 Sensitivity analysis with respect to gain splitting and the slope of the Engel curve

Our quantitative descriptions of the distributional effects of joint tax evasion depend on some fairly strong assumptions. To test how robust our results are with respect to the main assumptions of the framework, we consider

Table 7: Distribution of the tax compliance rate, 2009

Decile	Actual tax revenue	Full compliance	
		tax revenue	Compliance rate (%)
1	39,179	39,594	98.95
2	66,770	67,596	98.78
3	102,069	103,735	97.39
4	130,507	133,340	97.87
5	155,208	159,395	97.37
6	189,660	195,307	97.11
7	235,588	242,317	97.22
8	289,833	298,941	96.95
9	369,117	383,415	96.27
10	669,087	704,194	95.01
All	224,701	232,786	96.53

Notes: Households ranked by tax-evasion-adjusted post-tax income in 2009. All values refer to decile mean. Engel elast. and price assumption: $\varphi = 1$, $p_r = p_h$. Revenue effects from indirect and direct taxation are derived from tax-benefit model LOTTE (Aasness et al. 2007) and national account data (effective indirect tax rates). Marginal propensity to consume is set to 0.7 for all income groups.

alternative scenarios. As discussed in Section 5.1, we would like to see if the main results survive under different assumptions concerning the split of the economic gain between sellers and buyers and with respect to our assumption about the slope of the Engel curve (φ). Recall that the split of the economic gain is regulated by how the non-tax price of the hidden market, p_h , relates to the before tax price of the regular market, p_r . As the supplier gain is defined by the (hypothetical) income tax ($x_r^* \tau$) on non-reported income, if the price in the hidden market is low and falls short of the price in the regular market (before tax) ($p_h < p_r$), x_r^* decreases, and more of the gain is transferred to the buyers. Ultimately, the buyers may seize the whole advantage (or, vice versa, all of the gains may go to the suppliers).

As in the descriptions of distributional effects seen so far, we shall both describe effects in terms of supplier and consumer gains, SG and CG , and in terms of effects on the income distribution, in Table 8 and Table 9, respectively. As an alternative to $\varphi = 1$, we set $\varphi = 0$, i.e., assuming that the amount of hidden purchases are independent of income. However, note that the empirical approach still allows the probability of purchasing in the hidden market to depend on income. Further, in addition to the alternative used so far, which involves that consumer prices in the hidden market are 20 percent lower than in the regular market, corresponding to $p_h = p_r$, results

for 10 and 30 percent lower prices are presented in Table 8, whereas Table 9 also shows results for 0 percent and 40 percent lower prices in the hidden market.

The change to $\varphi = 0$ implies that the distributional gains of consumers are moved to lower parts of the income distribution. This can be seen by comparing CG of the middle alternative in Table 8 with the results reported in Table 5. When this is combined with a large price reduction in the hidden market (30 percent), the gains of consumers increase and, in particular, gains are large for the lower part of the distribution.

However, Table 9 shows that the income inequality (for all alternatives) is higher than the income inequality of conventional income, the latter reported in Table 4. The change to $\varphi = 0$ reduces the overall income inequality from 0.365 (see Table 6) to 0.363, and when more of the benefits are transferred to buyers, income inequality decreases further.

We see that all estimates of evasion-controlled income inequality are higher than the income inequality obtained from using standard income concepts. The clear pro-rich pattern of the suppliers is a main explanation. But, this conclusion cannot automatically be transferred to the case where we address the distribution of tax savings. Then the split of gains between suppliers and consumers becomes particularly decisive. If the consumers seize most to the gain, and the expenditure function is constant with respect to income ($\varphi = 0$), the sizeable gains of the consumers at the low end of the income distribution may dominate; see Table 8. However, for most alternatives of the tax savings perspective, we see distributional effects that point to tax evasion benefitting the rich more than the poor.

Table 8: Economic gains as share of tax evasion adjusted income under different consumer price reduction assumptions, $\varphi = 0$

Decile	Price reduction: 10%		Price reduction: 20%		Price reduction: 30%	
	<i>SG</i> (%)	<i>CG</i> (%)	<i>SG</i> (%)	<i>CG</i> (%)	<i>SG</i> (%)	<i>CG</i> (%)
1	0.28	0.64	0.23	1.44	0.17	2.44
2	0.28	0.37	0.25	0.83	0.21	1.42
3	0.55	0.34	0.50	0.76	0.43	1.30
4	0.76	0.33	0.70	0.74	0.63	1.26
5	0.92	0.33	0.85	0.75	0.75	1.28
6	1.01	0.34	0.86	0.76	0.67	1.30
7	0.82	0.34	0.70	0.76	0.54	1.30
8	0.93	0.33	0.79	0.74	0.62	1.26
9	1.40	0.31	1.20	0.69	0.95	1.19
10	2.82	0.22	2.44	0.50	1.96	0.86
All	1.40	0.31	1.22	0.70	0.98	1.19

Notes: Households sorted by tax-evasion-adjusted post-tax income in 2009.

All values refer to decile mean.

Table 9: Reported post-tax income as share of tax evasion adjusted post-tax income under different consumer price reduction assumptions, $\varphi = 0$

Decile	Lower price in the hidden market (%)				
	0	10	20	30	40
1	99.39	98.75	97.96	96.96	95.67
2	99.58	99.21	98.75	98.17	97.40
3	99.31	98.98	98.56	98.03	97.33
4	99.23	98.90	98.49	97.97	97.29
5	98.95	98.62	98.21	97.68	97.00
6	97.81	97.48	97.06	96.54	95.84
7	98.21	97.87	97.46	96.93	96.23
8	97.98	97.66	97.26	96.75	96.07
9	97.02	96.72	96.34	95.87	95.24
10	94.16	93.95	93.69	93.35	92.91
All	97.19	96.89	96.51	96.03	95.40
Gini coefficient, post-tax income	0.364	0.363	0.363	0.363	0.362

Notes: Households sorted by tax-evasion-adjusted post-tax income in 2009.

All values refer to decile mean.

6 Summary

The distributional effect of tax evasion in the informal economy depends on the income profile of both the buyers and sellers, as well as how the gain from evasion is shared between the two sides of the market. In this paper we offer a quantitative assessment of this question for Norway by evaluating data from both sides of the market. To learn about the suppliers of informal goods and services, we use the expenditure method. To evaluate the consumer side, we utilize sample survey information. Finally, we incorporate these two sources of information into a general equilibrium model that imposes consistency between supply and demand and which allows us to do robustness about the shifting parameter that we do not observe. This type of comprehensive analysis has not been attempted before for any country.

We find that accounting for the hidden economy in this comprehensive way increases the measure of income inequality in Norway, with the income profile distribution of informal suppliers contributing the larger share of the change in measured inequality. Moreover, the compliance rate of suppliers decreases with income, implying that the effective income tax progressivity is less than is indicated by official figures.

These conclusions must be provisional, as both the data and the methods of inference from the data are imperfect, and we offer these conclusions not as the final word on this issue. We do argue, though, our methodology show that this kind of comprehensive approach holds promise to shed light on the distributional effects of the informal economy.

References

- Aasness, J., J. K. Dagsvik and T.O. Thoresen (2007). The Norwegian Tax-benefit Model System LOTTE. In A. Gupta and A. Harding (Eds.): *Modelling Our Future: Population Ageing, Health and Aged Care*, International Symposia in Economic Theory and Econometrics, 513–518. Amsterdam: Elsevier.
- Abraham, M., K. Lorek, F. Richter, and M. Wrede (2015). *Collusive Tax Evasion and Social Norms*. CESifo Working Paper No. 5167, Munich.
- Allingham, M. G. and A. Sandmo (1972): *Income Tax Evasion: A Theoretical Analysis*, *Journal of Public Economics*, 1, 323–338.
- Alm, J. (2012). *Measuring, Explaining, and Controlling Tax Evasion: Lessons from Theory, Experiments, and Field Studies*. *International Tax and Public Finance*, 19, 54–77.

- Alm, J. and K. Finlay (2013). Who Benefits from Tax Evasion. *Economic Analysis and Policy*, 43, 139–154.
- Alm, J. and E. Sennoga (2010). Mobility, Competition, and the Distributional Effects of Tax Evasion. *National Tax Journal*, 63, 1055–1084.
- Andreoni, J., B. Erard, and J. Feinstein (1998). Tax Compliance. *Journal of Economic Literature*, 36, 818–60.
- Banks, J., R. Blundell, and A. Lewbel (1997). Quadratic Engel Curves and Consumer Demand. *Review of Economics and Statistics*, 79, 527–539.
- Benedek, D. and O. Lelkes (2011). The Distributional Implications of Income Under-Reporting in Hungary. *Fiscal Studies*, 32, 539–560.
- Bishop, J. A., J.P. Formby, and P. Lambert (2000). Redistribution through the Income Tax: The Vertical and Horizontal Effects of Noncompliance and Tax Evasion. *Public Finance Review*, 28, 335–350.
- Boadway, R., N. Marceau, and S. Mongrain (2002). Joint Tax Evasion. *Canadian Journal of Economics*, 35, 417–35.
- Chang, J.-j. and C.-c. Lai (2004). Collaborative Tax Evasion and Social Norms: Why Deterrence Does Not Work. *Oxford Economic Papers*, 56, 344–368.
- Christian, C.W. (1994). Voluntary Compliance with the Individual Income Tax: Results from the 1988 TCMP Study, *IRS Research Bulletin*, 1993/1994, Publication 1500, 35–42.
- Cremer, H. and F. Gahvari (1993). Tax Evasion and Optimal Commodity Taxation. *Journal of Public Economics*, 50, 261–275.
- De Økonomiske Råd (2011). *Dansk Økonomi, efterår 2011*, Copenhagen (in Danish, with English summary).
- Dhami S. and A. al-Nowaihi (2007). Why Do People Pay Taxes? Prospect Theory versus Expected Utility Theory. *Journal of Economic Behavior and Organization*, 64, 171–192.
- Dunbar, G.R., and C. Fu (2015). Sheltered Income: Estimating Income Under-Reporting in Canada, 1998 and 2004. *Bank of Canada Working Paper No. 2015-22*, Ottawa.

- Eide, E., K. von Simson, and S. Strøm (2011). Rank-Dependent Utility, Tax Evasion and Labor Supply. *FinanzArchiv/Public Finance Analysis*, 67, 261–281.
- Engström, P. and B. Holmlund (2009). Tax Evasion and Self-Employment in a High-Tax Country: Evidence from Sweden, *Applied Economics*, 41, 2419–2430.
- Engström, P. and J. Hagen (2015). Income Underreporting among the Self-Employed: A Permanent Income Approach. Working Paper 2015:2, Uppsala Center for Fiscal Studies, Department of Economics, University of Uppsala.
- Feldman, N.E. and J. Slemrod (2007). Estimating Tax Noncompliance from Unaudited Tax Returns. *Economic Journal*, 117, 327–352.
- Feldstein, M. (1972). Distributional Equity and the Optimal Structure of Public Prices. *American Economic Review*, 62, 32–6.
- Fiorio, C.V. and F. D’Amuri (2005). Workers’ Tax Evasion in Italy. *Giornale degli Economisti e Annali di Economia*, 64, 241–264.
- Fortin, B., G. Lacroix, and D. Pinard (2010). Evaluation of the Underground Economy in Quebec: A Microeconomic Approach, *International Economic Journal*, 24, 463–479.
- Groves, R.M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, 70, 646–675.
- Holmøy, A. and M. Lillegård (2014). Forbruksundersøkelsen 2012. Dokumentasjonsrapport. Notater 2014/17, Statistics Norway (In Norwegian).
- Hurst, E., G. Li, and B. Pugsley (2014). Are Household Surveys Like Tax Forms: Evidence from Income Underreporting of the Self Employed. *Review of Economics and Statistics*, 96, 19–33.
- Johansson, E. (2005). An Estimate of Self-Employment Income Underreporting in Finland, *Nordic Journal of Political Economy*, 31, 99–109.
- Johns, A. and J. Slemrod (2010). The Distribution of Income Tax Noncompliance, *National Tax Journal*, 63, 397–418.
- Kesselman, J.R. (1989). Income Tax Evasion: An Intersectoral Analysis. *Journal of Public Economics*, 38, 137–182.

- Kim, B., J. Gibson, and C. Chung (2016). Using Panel Data to Estimate Income Under-Reporting by the Self-Employed. The Manchester School, forthcoming.
- Kleven, H.J., M. Knudsen, C.T. Kreiner, S.L. Pedersen, and E. Saez (2011). Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark, *Econometrica*, 79, 651–692.
- Kleven, H.J., C.T. Kreiner, and E. Saez (2016). Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries, forthcoming, *Economica*.
- KRISINO (2011). Kriminalitets- og sikkerhetsundersøkelsen i Norge. Næringslivets sikkerhetsråd (In Norwegian).
- Lyssiotou, P., Pashardes, P. and Stengos, T. (2004). Estimates of the Black Economy Based on Consumer Demand Approaches, *Economic Journal*, 114, 622–640.
- Martinez-Lopez, D. (2015). The Underreporting of Income by Self-Employed Workers in Spain. *SERIEs*, 4, 353–371.
- Matsaganis M. and M. Flevotomou (2010). Distributional Implications of Tax Evasion in Greece, GreeSE Paper No 31, Hellenic Observatory Papers on Greece and Southeast Europe, LSE, London.
- Ognedal, T. (2015). Morality in the Market. Manuscript, University of Oslo.
- Opinion (2009). Privates kjøp av svart arbeid, Skattedirektoratet. (in Norwegian)
- Pashardes, P. and A. Polycarpoub (2008). Income Tax Evasion, Inequality and Poverty. *Cyprus Economic Policy Review*, 2, 37–49.
- Paulus, A. (2015a). Tax Evasion and Measurement Error: An Econometric Analysis of Income Survey Data Linked with Tax Records. ISER Working Paper 2015-10, University of Essex, Colchester.
- Paulus, A. (2015b). Income Underreporting Based on Income-Expenditure Gaps: Survey vs Tax Records, ISER Working Paper Series, No. 2015-15, University of Essex, Colchester.
- Persson M. and P. Wissén (1984). Redistributive Aspects of Tax Evasion. *The Scandinavian Journal of Economics*, 86, 131–149.

- Pissarides, C.A. and G. Weber (1989). An Expenditure-Based Estimate of Britain's Black Economy, *Journal of Public Economics*, 39, 17–32.
- Schneider, F. and D.H. Enste (2000). Shadow Economies: Size, Causes, and Consequences. *The Journal of Economic Literature*, 38, 77–114.
- Schuetze, H. (2002). Profiles of Tax Non-Compliance among the Self-Employed in Canada: 1969–1992, *Canadian Public Policy*, 28, 219–238.
- Slemrod, J. (2007). Cheating Ourselves: The Economics of Tax Evasion, *Journal of Economic Perspectives*, 21, 25–48.
- Slemrod, J. and S. Yitzhaki (2002). Tax Avoidance, Evasion, and Administration. In A.J. Auerbach and M.S. Feldstein (Eds.): *Handbook of Public Economics*, Volume 3, 1423–70. Amsterdam: Elsevier.
- Slemrod, J. and C. Weber (2012). Evidence of the Invisible: Toward a Credibility Revolution in the Empirical Analysis of Tax Evasion and the Informal Economy. *International Tax and Public Finance*, 19, 25–53.
- Statistics Norway (2005). *Income Statistics for Persons and Families 2002–2003*, Official Statistics of Norway.
- Tedds, L.M. (2010). Estimating the Income Reporting Function for the Self-Employed, *Empirical Economics*, 38, 669–687.
- TNS Gallup (2006). *Kjøp av svart arbeid i privatmarkedet*, Skattedirektorat (in Norwegian).
- Torgler, B. (2007). *Tax Compliance and Tax Morale*. Cheltenham: Edward Elgar.
- Wangen, K.R. (2005). An Expenditure Estimation of Britain's Black Economy Revisited, Discussion Papers, No. 414, Statistics Norway.
- Williams, C.C. (2008). *The Hidden Enterprise Culture: Entrepreneurship in the Underground Economy*. Cheltenham: Edward Elgar.
- Yaniv, G. (1988). Withholding and Non-Withheld Tax Evasion. *Journal of Public Economics* 35, 183–204.
- Yaniv, G. (1992). Collaborated Employee-Employer Tax Evasion. *Public Finance*, 47, 312–21.
- Yitzhaki, S. (1974). A Note on "Income Tax Evasion: A Theoretical Analysis", *Journal of Public Economics*, 3, 201–202.

A Data description

Data from Statistics Norway's Survey of Consumer Expenditures over the period from 2003 to 2009 are used in the expenditure approach, described in Section 3. Each year 2200 persons between 0–79 have been randomly drawn from the Norwegian population. The average response rate is below 50 percent, which means that we have access to a net sample of approximately 1000 households each year. The survey is based on personal interviews and detailed accounting of household expenditures for a period of 14 days. In addition to expenditures, the data include household characteristics, some of them (as income and education) by linking to administrative registers.

As described in Section 4, we employ measures of permanent income, based on 7-year averages, in some of the estimations. For a given survey year, we add income of 3 preceding years and the 3 subsequent years to reach a 7-year average. As we cannot merge this additional income information to the expenditure data by personal ID numbers, the linking is done by exploiting that there is overlapping information in the two data sets. In addition to net household income, we use a regional variable, number of household members, gender composition and age group of the oldest member to merge datasets. By using this approach we are not able to get an unique match on all households, but we find match for around 70 percent of the households.

Further, in line with earlier contributions using the expenditure approach, we exclude the self-employed of the primary sectors. We also restrict the sample to households with net permanent income below 2.5 million NOK. In the upper part of Table A1 we report descriptive statistics for the variables used in the estimation of tax evasion among suppliers. We define a household as a "self-employment household" if it contains at least one member with the self-reported employment status as self-employed. Moreover, a household is defined as a "wage earner household with potential for tax evading" if all members are wage earners and at least one householder member reporting "craftsman" as the occupation.

As seen in Section 4, the demand side behavior is described by using information from two sample surveys that were carried out in 2006 and 2009; see TNS Gallup (2006) and Opinion (2009). The response ratio is 0.56 in the 2009 survey and most likely somewhat lower in 2006. We pool the information and end up with more than 3,000 respondents, approximately 1,000 from the 2006-survey and 2,000 from the 2009-survey. In the surveys respondents are asked about their participation in hidden market transactions, and how much they have bought during the two last years. In the estimation of the probit equation we use information about households with less than 3.5 million NOK

Table A1: Descriptive statistics

Supply side (expenditure approach)						
	Wage earners		Self-employed		Craftsmen	
	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.
Log net income	13.02	.52	13.09	.56	13.15	.35
Log perm. net income	13.09	.48	13.17	.48	13.19	.34
Log food expenditures	10.71	.67	10.89	.60	10.87	.53
Age	49.11	13.33	47.79	10.35	44.47	9.63
Number of adults	1.83	.55	1.94	.46	2.11	.57
Number of children	1.18	1.24	1.44	1.23	1.52	1.24
House size (m^2)	139.23	61.27	157.57	68.16	151.90	60.62
High education dummy	.34	.47	.30	.46	.11	.31
Log capital income	7.90	2.43	8.32	2.24	7.67	2.25
Observations	3494		252		467	
Demand side (probit estimation)						
	Mean		Std.dev.			
Binary for purchases	.15		.35			
Log gross income	13.20		.57			
Age	48.06		14.57			
Male	.53		.50			
High education dummy	.29		.45			
Self-employed	.05		.22			
Eastern region dummy	.17		.37			
Observations			3259			

Note: In the supply side sample age refer to age of the main income earner of the household.

in income. Descriptive statistics for the variables used in the estimation are reported in the lower part of Table A1.

Recall that we do not use information about evaded amounts from this data source, only whether the individual has been involved in tax evasion or not. Instead we proceed by specifying demand functions and calibrate demand parameters according to what we perceive as the best available knowledge at the moment, which includes results found in De Økonomiske Råd (2011), and employ information from the supply side when we calibrate total amounts evaded.

However, Table A2 reports the evaded amounts as seen in TNS Gallup (2006) and Opinion (2009), although in very wide intervals, which is the main reason for these data being less useful. A clear majority of those who

Table A2: Number of observations sorted by annual gross household income and hidden market expenditures. Pooled information, 2006 survey and 2009 survey

Income (NOK)	Expenditures (NOK)					All
	0	0–10K	10K–50K	50K–100K	>100K	
0–200,000	148	7	1	0	0	156
200K–500K	1163	114	44	4	1	1326
500K–1M	1255	165	61	7	4	1492
1M–3M	202	32	25	3	0	262
>3M	168	11	8	0	1	188
All	2936	329	139	14	6	3424

Notes: Expenditures measured over 2 years.

1€ = 8.73 Norwegian kroner (NOK), and 1\$ = 6.64 NOK in 2009.

buy hidden services reports to have purchased for less than 50,000 NOK. Also, evaded amount appears to increase with gross income: among evaders, the share spending more than 10,000 NOK increases with the income level (except in the highest income group).