Distributional Implications of Joint Tax Evasion*

Odd E. Nygård, Joel Slemrod and Thor O. Thoresen†

Abstract

Both buyers and sellers of goods and services 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, and therefore a lower priced product. The implications of such joint tax evasion for income distribution 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 magnitudes and the distribution of evaded income among sellers and buyers of goods and services. Our preliminary results indicate that the tax-evasion-controlled estimate of income inequality in Norway exhibits more income dispersion than official estimates.

1 Introduction

Who gain most from tax evasion, the rich or the poor? This is a complicated question as tax evasion induces numerous effects on factor and commodity

---

*Financial support from the Norwegian Research Council is gratefully acknowledged.
†Nygård: Research Department, Statistics Norway, Postboks 8131 Dep. NO-0033 Oslo, Norway (e-mail: oen@ssb.no); Slemrod: Stephen M. Ross School of Business and Department of Economics, University of Michigan, 701 Tappan Street, Ann Arbor, MI 48109 (e-mail: jslemrod@umich.edu); Thoresen: Research Department, Statistics Norway, Postboks 8131 Dep. NO-0033 Oslo, Norway, and Oslo Fiscal Studies, University of Oslo (e-mail: tot@ssb.no).
prices (Kesselman, 1989; Slemrod, 2007; Alm and Finlay, 2013). Even 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 parameterize 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. 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. The seller can then evade income tax and the buyer benefits from paying a lower price for the commodity or services, also because there is no VAT remittance involved. The overall distributional consequences of including the economic gain of hidden transactions depend on how sellers and buyers divide the economic advantage between them, and it depends on how both the sellers and buyers are positioned in the distribution of income.

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 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 transaction perspective. By piecing together empirical evidence obtained from the supply side and the demand side of the market, income inequality is calculated accounting for the underreporting of income among the suppliers of services and after adjusting the distribution of disposable income for the differences between buying services in the legal versus illegal part of the market. One objective is to present a more accurate picture of the distribution of economic well-being in the society, compared to official figures of income inequality.

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 "indirect methods" 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 methods approach, evasion is not measured directly, but indirectly via measurable traces. Given that several groups of wage earners have rather limited scope for tax evasion (because third-party reporting of income is a standard proce-
dure), 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 calculate the amount by which reported income must be scaled up in order to obtain true income levels for tax evaders.\footnote{The so-called "expenditure approach", set forth by Pissarides and Weber (1989) and exemplified by an application on British data, have 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, Martínez-Lopez (2013) for Spain, Paulus (2015) for Estonia, and Feldman and Slemrod (2007) and Hurst, Li and Pugsley (2014) for the U.S.} We allow, to some extent, for variations in the degree of underreporting along the income range.

With respect to the buyers’ side, we rely on "direct methods" for measurement of tax evasion, in that survey data about purchases of illegal services in Norway are exploited (TNS Gallup, 2006; Opinion, 2009). We obtain estimates of the probability of being involved in transactions not reported to the tax authorities and the amount spent, depending on different households characteristics, such as income and education.

The parameter estimates from the two econometric approaches are used to predict individual income gains which generate an estimated effect on the distribution of income. 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. Thus, we are able to discuss how the "hidden-economy-controlled" income distribution compares to the official one: is it less or more equal?

The paper is organized as follows. Section 2 summarizes some of the main perspectives and its distributional effects. 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 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. Similarly, 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 the decision-making. One key criticism is that taxpayers are not motivated by narrow self-interest alone, but act as a member of a group, which means that taxpayers also are influenced by norms, customs, reciprocity, patriotism, etc. Further, individuals are obviously not only motivated by economic factors; notions such as shame, guilt and morality arguably also influence decisions in some circumstances. 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, and we will use it as the theoretical foundation of our inquiry.

The present study challenges the predominant perspective in the literature that tax evasion occurs as interaction between a single economic agent and the government. Previous studies have elaborated on collusive tax cheating between employees and the employer, see Yaniv (1988; 1992) and Kleven et al. (2014). Further, 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).

When turning attention towards studies on distributional gains of tax evasion, the market transaction perspective of the present study implies more 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 one can in principle control for a whole range of reactions by individuals and firms. Sandmo (1981) discusses welfare implications of choices of tax rates, penalties and detection within an optimal tax model, whereas Persson and Wissen (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, and in such a model Alm and Sennoga (2010) 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 in the present context. Bishop, Formby, and Lambert (2000) and Johns and Slemrod (2010) use data from the auditing systems of the Internal Revenue Service (IRS) of the US (the Taxpayer Compliance Measurement Program, and the National Research Program, respectively), which allow the researchers to observe income as reported and as adjusted by an audit. Bishop et al. find that, for the tax year 2001, 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 small impact on the Gini coefficient. Johns and Slemrod (2010) find that tax noncompliance make the true income distribution more unequal than what we observe, but the tax system becomes more progressive. This follows from a given percentage reduction in taxable income corresponding to a particularly high percentage reduction in tax liability for taxpayers with taxable income just above the taxpaying threshold.\(^3\) Kleven et el. (2011)

\(^2\)See also the discussion in Alm and Finlay (2013).
\(^3\)Christian (1994) also analyzes data from the Taxpayer Compliance Measurement Program to discuss distributional aspects of tax evasion. It suggests 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.
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 evade more as a share of self-reported income than those with relatively high self-reported income.

Pashardes and Polycarpou (2008) employ an expenditure approach technique, outlined in Lyssiotou, Pashardes and Stengos (2004), 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 finds that the GST increased tax noncompliance by those with larger amounts of self-employment income and unaffected tax noncompliance by those with small amounts of self-employment income.

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 might consider declaring a higher figure to an anonymous interviewer. Fiorio and D’Amuri (2005), Matsaganis and Flevotomou (2010) and Benedek and Orsolya (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 fell with income, Matsaganis and Flevotomou (2010) suggest tax evasion gives higher income inequality and more poverty and lower progressivity of the income tax, which is also in line with findings in Benedek and Orsolya (2011).

3 Theoretical framework

We assume there are two commodities in the economy, an 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. Subscripts $i$ are suppressed in what follows.

\footnote{Instead of using expenditures on food only, as in Pissarides and Weber (1989), Lyssiotou et al. use information on a whole range of consumer goods.}
3.1 Supply side

Each supplier has a skill level denoted $n$, and a skill type, $\varphi_{c1}$ and $\varphi_{c2}$, 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 $\varphi_{c2}$, and zero otherwise. A supplier of type $\varphi_{c2}$ decides whether to report the income for tax purposes, remit tax at rate $\tau$, and thereby supply the service formally. Alternatively, he does not report the income and supplies the service in the informal market. For simplicity, we assume that each supplier reports all or none of its income from producing $c_2$. He makes this decision by comparing the expected utility from underreporting to the utility from reporting truthfully. In the case when the tax evasion is discovered, which occurs with probability $\rho$, the evader has to pay a penalty, $\theta$, in addition to the evaded tax. In addition, tax evasion incurs a psychic costs $\Psi$, which varies across individuals.

Let $p_r$ be the net price (before-sales-tax price) in the regular market, and $p_h$ the price in the informal market. Further, we let $l_h$ and $l_r$ denote the optimally chosen labor supply in each sector for the supply of $c_2$. Then we let $x_h$ and $x_r$ equal true before-tax income if operating in the regular or hidden economy, where $x_h = np_h l_h$ and $x_r = np_r l_r$. Thus, a supplier of $c_1$ enters the hidden economy if

$$\psi((1 - \theta x_h) - \theta x_h) - \Psi > x_r(1 - \tau)$$

i.e. if the expected utility of entering the hidden market is larger than receiving the certain after tax income $x_r(1 - \tau)$ in the regular market.

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 $\varphi_{c2}$ change when the skill level increases. We also define

$$k = \frac{y^*}{y}.$$  \hspace{1cm} (2)

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 $d_{yk}$ as the income elasticity characterizing the change in $k$ when income $y$ increases. These elasticities, like Feldstein’s "distributional characteristics" (Feldstein, 1975), 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 $d_{ns} > 0$ the probability of being of type $\varphi_{c2}$ decreases with the skill level. Hence, we will find more of this type at the bottom of the income

---

5It is assumed that both the seller and buyer know with certainty whether a transaction will be reported for tax purposes.
distribution and adjusting the income distribution for unreported income could tend to make it more equal, and ii) if $d_{yk} < 1$ the underreported income as share of the total income decreases with the reported income, thus adjusting for unreported income makes the distribution more equal, ceteris paribus.

The next question to address is how much they gain from tax evasion. If we assume that total labor supply is fixed ($l_h = l_R$), and the price in the informal market is equal to the pre-sales tax price in the regular ($p_h = p_r$), the suppliers will save the income tax by entering the informal market instead of supplying in the regular market. When adding the risk of being caught, their expected earning is $(1 - \rho)x\tau - \rho\theta x = x\tau + x\rho(\tau + \theta)$. Note that if the probability of getting caught is low, this approximately equals the saved income tax. How much income tax they save depends on the shape of the income tax scheme, e.g. high-income individuals may benefit more if the tax system is progressive, i.e. the tax evasion makes the tax system less progressive. Without any psychic costs, this will coincide with the compensated variation: what we need to pay the supplier in the hidden market to be equally well off if entering the regular market. But, since there is some psychic costs associated with supplying in the hidden market, the actual compensation needed will be somewhat lower. In addition, attitudes towards risk will come more into play when the probability of being caught increases. Finally, labor supply varies with prices, and the suppliers may want to change their labor supply if operating in the hidden market, this will also tend to overstate the compensation needed. However, taking into account such effects require knowledge about preferences and psychic costs, so we will refrain from focusing on this and instead compute only the pecuniary gains.

3.2 Demand side

Each consumer demands 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 amounts denoted $c_h$ and $c_r$. If he purchases in the hidden market and gets caught, he will get penalized by a certain fraction of the tax evaded, denoted $\kappa$. In addition, there is psychic cost, $\chi$, which may vary between individuals, presuming that the consumer is aware of whether the supplier is tax-compliant. Thus, if entering the hidden market he is faced with the decision problem

$$\max_u u(c_1, c_2, l) - \chi \ s.t. \ y^* = c_1 + p_h c_h,$$

---

6 This will affect the shape if the $u(.)$ function.
if not caught, and

\[
\max u(c_1, c_2, l) - \chi \quad \text{s.t.} \quad y^* = c_1 + p_h (1 + \pi) c_h + \kappa p_h \pi c_h, \quad (4)
\]

if caught and penalized, where \( \pi \) is the sales tax. Note that because the consumer could also be a supplier, conditional on the skill level and the prices, \( y^* \) will depend on the amount of labor he supplies and the income tax he pays.\(^7\) Define the corresponding indirect utility functions as

\[
v(n, p_h, \chi), \quad (5)
\]

\[
v(n, p_h, \pi, \kappa, \chi). \quad (6)
\]

If we let \( \eta \) denote the probability of getting caught tax and being penalized, we can summarize this problem by defining

\[
V_h(n, p_h, \pi, \kappa, \chi, \eta) \equiv (1 - \eta)v(n, p_h, \chi) + \eta v(n, p_h, \pi, \kappa, \chi), \quad (7)
\]

i.e. the indirect expected utility of entering the hidden market as a function of the different parameters. The consumer buys in the hidden market if

\[
V_h(n, p_h, \pi, \kappa, \chi, \eta) > V_r(n, p_r(1 + \pi)), \quad (8)
\]

i.e. the expected utility of the uncertain basket from the hidden-market choice exceeds the utility from the certain consumption basket from the regular-market decision.

Supposing that \( p_r = p_h \), \( c_r = c_h \) and labor supply is fixed, the gain of entering this market is the expected money saved which is equal to \((1 - \eta) p_h \pi c_h - \eta \kappa p_h \pi c_h\). Next, the gain as share of the total income can be written as

\[
b = \frac{p_h \pi c_h (1 - \eta(1 + \kappa))}{Y^*}, \quad (9)
\]

and we define an elasticity \( d_{Y \cdot b} \) describing how the budget share, \( b \), of hidden market gains change when income changes. If \( d_{Y \cdot b} > 1 \), the share increases with income, and purchases in the hidden market makes the income distribution more unequal. This will be the case if the demand for the hidden good is of a luxury type, with an Engel elasticity above one. Note that if the probability of getting caught is low, we can approximate the economic gain to be the saved sales tax. As above, this economic gain will deviate somewhat to the compensated variation due to risk, psychic costs and the fact that

\(^7\)We assume for simplicity that if he gets caught, he must pay by decreasing the demand for the numeraire commodity or increasing labor supply.
the consumer could adjust his consumption bundle and labor supply when entering the regular sector. In addition we also have the issue of splitting the tax saving between the supply and demand side. However, assuming that the gain is the saved sales tax, we can write the expected gains (as share of income) for a certain individual as

\[ \Pr \left( c_h > 0 \mid Z' \right) \frac{p_h \pi c_h}{Y_s} = \Pr \left( c_h = 1 \mid Z' \right) b, \]

(10)

where \( \Pr \left( c_h > 0 \mid Z' \right) \) is the probability of the individual entering the hidden market, depending on characteristics \( Z' \). Let the elasticity \( d_Y \cdot \Pr \) define how the probability of entering the market depends on income. Then, the distributional effects will depend on both \( d_Y \cdot b \) and \( d_Y \cdot \Pr \), i.e. how likely he is to enter given his income, and how much he will purchase.

3.3 Equilibrium

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

\[ F_h(p_{y_r}, p_{y_h}, \tau, \theta, \rho) = D_h(p_{y_r}, p_{y_h}, \pi, \kappa, \eta), \]

(11)

and

\[ F_r(p_{y_r}, p_{y_h}, \tau, \theta, \rho) = D_r(p_{y_r}, p_{y_h}, \pi, \kappa, \eta), \]

(12)

where \( F(.) \) and \( D(.) \) are the aggregate supply and demand, respectively, and subscript denotes the market. In particular, the shape of these functions will depend on the distribution of skills and the psychic costs, \( \Psi \) and \( \chi \), in the economy. If the probability of being caught and the psychic costs are low for the whole population, suppliers are easily attracted to the hidden sector and prices are driven down such that the after-tax income of supplying a certain amount approximates the income from selling in the hidden market. In this case there will be nearly no economic gains for the supplier of being in the hidden sector, while all the gains are received by the consumers.

On the other hand, if there are in general large psychic costs, large compensation is needed, and the hidden market price will tend to be higher. The steepness of the informal market supply curve depends on the distribution of \( \Psi \) in the economy. As long as the supply curve is not perfectly elastic, the position of the demand curve will then determine the market equilibrium price.

Note that in equilibrium the following must be true...
\[ \sum x_h = \sum p_h c_h, \]  
\[ (13) \]
i.e. the sum of evaded incomes equals the sum of hidden payments from the consumers.\(^8\)

4 Quantifying tax evasion at both sides of the market

4.1 Identification of evaded income of suppliers

The expenditure approach follows from the assumptions that some individuals have the opportunity to underreport, referred to as skill type \( \varphi_{c2} \) in Section 3.1, while others do not, and that the groups have similar preferences for a consumption good. Thus, it is assumed that for everyone consumption, \( c \), is determined by true, permanent income, \( y^* \), and a number of individual control variables, \( Z' \). Thus, when using the log form, we have the following Engel curve relationship, \( \ln c = Z' \gamma + \beta \ln y^* \), where \( \gamma \) and \( \beta \) are parameters.\(^9\) Pissarides and Weber (1989) let \( c \) be represented by consumption of food and assume that the self-employed is the group which has scope for underreporting.

As already seen in Equation (2), the relationship between observed income, \( y_{ij} \), and true, permanent income, \( y_{ij}^* \), for individual \( i \) of type \( j \) can be described by a proportionality factor, \( k_{ij} \), where we assign type \( j \) to indicate that there are differences across individuals in the scope for evasion. Thus, \( k_{ij} (> 1) \), shows the factor by which the observed income for individual \( i \) of type \( j \) can be multiplied in order to become equal to the true income. Influenced by Pissarides and Weber (1989) and several subsequent papers adopting their line of research, we let one of the groups (\( j \)) be represented by the self-employed. However, we shall allow for some groups of wage earners to be 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.\(^10\)

Further, following Pissarides and Weber, standard applications of the expenditure approach assume that observable income fluctuates around per-

\(^8\)We suppose that they are subject to regular prices and taxes on factor of inputs.
\(^9\)Thus, 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).
\(^10\)For 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.
manent income by a factor \( g \), seen as \( y_{ij} = g y_{ij}^* \), and usually assume that the coefficients \( \ln g_{ij} \) and \( \ln k_{ij} \) and are lognormally distributed around their means, \( \ln g_{ij} = \mu_k + u_k \) and \( \ln k_{ij} = \mu_k + v_{ij} \). Then it follows that the relationship between (true) permanent income and observable income is seen as \( \ln y^* = y_{ij} - (\mu_k - \mu_k) - (u_{ij} - v_{ij}) \), and we have

\[
\ln c_{ij} = Z_{ij} \gamma + \beta \ln y_{ij} - \beta (\mu_k - \mu_k) + \beta (u_{ij} - v_{ij}) + \varepsilon_{ij}. \tag{14}
\]

If we focus on tax evasion among the self-employed, the Engel curve by adjusted by an indicator variable, \( q_i \), which takes the value 1 for the self-employed, \( SE \), and let the (non-evading) salary workers be symbolized by \( SW \), Equation (14) can be seen as

\[
\ln c_{ij} = Z_{ij} \gamma + \beta \ln y_{ij} + \beta (\mu_{kSW} - \mu_{gSW}) + \beta q_i \left[ (\mu_{kSE} - \mu_{kSW}) - (\mu_{gSW} - \mu_{gSE}) \right] + \beta (u_{ij} - v_{ij}) + \varepsilon_{ij}. \tag{15}
\]

Further, the mean of \( k_{SE} \) is given by \( \ln \kappa_{SE} = \mu_{kSE} + \frac{1}{2} \sigma^2_{vSE} \), where \( \sigma^2_{vSE} \) is the variance of \( v_{ij} \) for \( j \in SE \). Also, as \( k_{SW} = 1 \) for \( j \in SW \), and as the income means in the two groups are identical, \( \ln \overline{g}_{SE} = \ln \overline{g}_{SW} \), Equation (15) can be rearranged into the following reduced form, which is the standard empirical specification used to obtain estimates of \( k \),

\[
\ln c_i = Z_i \gamma + \beta \ln y_i + \delta q_i + \eta_i. \tag{16}
\]

As \( \delta = \beta \left[ \mu_{kSE} + \frac{1}{2} (\sigma^2_{vSE} - \sigma^2_{uSW}) \right] \), and \( \eta_i = \beta (u_{ij} - v_{ij}) + \varepsilon_{ij} \), an estimate of the adjustment factor \( k \) is given by

\[
\kappa_{SE} = \exp \left[ \mu_{kSE} + \frac{1}{2} \sigma^2_{vSE} \right] = \exp \left[ \frac{\delta}{\beta} + \frac{1}{2} (\sigma^2_{vSE} + \sigma^2_{uSW} - \sigma^2_{uSE}) \right]. \tag{17}
\]

However, as \( \sigma^2_{vSE} \), \( \sigma^2_{uSW} \), and \( \sigma^2_{uSE} \) are usually not known, a standard empirical approach along the lines of Pissarides and Weber involves obtaining estimates of the variance of the residuals, \( \zeta_{ij} \), from an expression, \( \ln y_{ij} = B_{ij} \psi + \zeta_{ij} \), where \( B_{ij} \) includes a set of instruments for permanent income. Thus, an estimate of \( \kappa_{SE} \) is obtained by\(^{11}\)

\[
\kappa_{SE} = \exp \left[ \frac{\delta}{\beta} \pm \frac{1}{2} (\sigma^2_{\zeta SE} - \sigma^2_{\zeta SW}) \right]. \tag{18}
\]

\(^{11}\)Pissarides and Weber discuss result for both a lower bound case (\( \sigma^2_{vSE} = 0 \)), and an upper bound alternative (\( \sigma^2_{uSW} = \sigma^2_{uSE} \)). See also Wangen (2005) on this.
The approach to obtain estimates of $k$ for specific groups of salary workers will follow the same type of reasoning. In our main specification we estimate Equation (16) by directly using a measure of permanent income for $y_i$, which simplifies Equation (18) somewhat.

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 $\delta$. This can very straightforwardly be done by introducing a dummy variable reflecting high income, $hi_i$, and interacting it with the variable exhibiting that there is scope for underreporting, $q_i$, in Equation (16).

\[
\ln c_i = \alpha + Z_i \gamma + \beta \ln y_i + \delta q_i + \lambda (q_i \times hi_i) + \eta_i. \tag{19}
\]

### 4.2 Data and estimation results for expenditure approach

Estimates of $k$ are obtained by employing data from the Norwegian Survey of Consumer Expenditure (Holmøy and Lillegård, 2014). As the sample sizes are small, we pool information over the time period from 2003 to 2009. Further, in order to obtain estimates of permanent income, we link these data to income panel data 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, we lose a small number of units. The measure of permanent income is obtained by taking averages over the most recent 7 years of income data. More information about the data and the procedure is provided in Appendix A.

The 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 wage earners with some scope for tax evasion, based on the specification in Equation (16). Results for specifications letting income be represented by both yearly income and permanent income are provided, using both OLS and instrumental variables techniques in the estimations. Given that our 7-year income average represents permanent income adequately, Equation (18) is simplified, as $\sigma^2_{uSW} = \sigma^2_{uSE}$ that the remaining contribution comes from variance in the self-employment underreporting rate ($\sigma^2_{vSE}$), and $\bar{k}_{SE} = \exp \left[ \frac{\delta}{\beta} + \frac{1}{2} \left( \sigma^2_{\zeta SE} - \sigma^2_{\zeta SW} \right) \right]$ is used to obtain estimates of $\bar{k}_{SE}$, see also Kim, Gibson and Chung (2009). As in most of the previous literature, including Pissarides and Weber (1989), estimates are also obtained by using IV methods. The results of the 2SLS estimations
are based on using size of house, dummy for higher education and (log of) capital income as instruments, and the F-statistic, Sargan’s overidentification test and Wu-Hausman endogeneity test are employed for diagnostics.\footnote{The choice of instruments is influenced by earlier studies, although there is substantial variation across studies (Paulus, 2015). Engström and Hagen (2015) report results that support using capital income as instrument.}

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 (19).

Estimates of $k$ show little variation depending on the choice of income measure and the estimation procedure. Estimates for the self-employed are found in the range from 1.20 to 1.25, and when using the preferable permanent income specification, estimates of $k_{SE}$ are 1.23 and 1.21.\footnote{In 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).} Thus, the OLS and IV estimates are relatively close.\footnote{Note that the F-statistic of the first stage, p-values of the Sargan’s overidentification test and Wu-Hausman endogeneity test all suggest good performance of the IV approach.}

These estimates are also not far from the estimates for the share of underreported income among the self-employed in Denmark (Kleven et al., 2011) and in Sweden (Engström and Hagen, 2015).\footnote{Note that both Kleven et al. (2011) and Engström and Hagen (2015) report results in terms of the fraction of income, which means that a recalculation is needed in order to compare our estimates for $k$.}

We see no signs of tax evasion among wage earners on average. However, when allowing for different effects by income range, we see clear indications of tax evasion among salary workers with income below the median, of between 8 and 10 percent. Also, when we differentiate between self-employed with high and low income, results suggest that there is more tax evasion at low (reported) income levels, although non-significantly for the permanent income specification only.

### 4.3 Buyers in the hidden market

When moving to evidence on tax evasion of the demand side, Section 3.2 describes the decision problem of purchasing agents who have the possibility to be involved in non-reported transactions. To obtain information about the buyers we exploit information from two sample surveys that were carried out in 2006 and 2009 to obtain information about the extent of the informal economy, presented in TNS Gallup (2006) and in Opinion (2009). Of course, given the topic of the surveys, there is a clear potential for nonresponse bias.
Table 1: Estimation results for expenditure approach. Pooled data, 2003–2009

<table>
<thead>
<tr>
<th></th>
<th>OLS, yearly</th>
<th>IV, yearly</th>
<th>OLS, permanent</th>
<th>IV, permanent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope, $\beta$</td>
<td>.322 (13.85)**</td>
<td>.332 (7.79)**</td>
<td>.401 (15.96)**</td>
<td>.374 (8.19)**</td>
</tr>
<tr>
<td>Self-employed, $\delta_{SE}$</td>
<td>.070 (2.27)**</td>
<td>.060 (1.92)*</td>
<td>.067 (2.16)**</td>
<td>.056 (1.77)*</td>
</tr>
<tr>
<td>Salary worker, $\delta_{SW}$</td>
<td>-.015 (-0.67)</td>
<td>-.018 (-0.80)</td>
<td>-.004 (-0.18)</td>
<td>-.010 (-0.45)</td>
</tr>
<tr>
<td>Implied $k_{SE}$</td>
<td>1.25 (4.20)**</td>
<td>1.20 (3.06)*</td>
<td>1.23 (5.89)**</td>
<td>1.21 (4.14)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.445</td>
<td>.446</td>
<td>.450</td>
<td>.452</td>
</tr>
<tr>
<td>1st stage F-stat</td>
<td>263.8</td>
<td>337.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sargan ($p$–value)</td>
<td>.230</td>
<td></td>
<td>.632</td>
<td></td>
</tr>
<tr>
<td>Wu-H ($p$–value)</td>
<td>.865</td>
<td></td>
<td>.344</td>
<td></td>
</tr>
</tbody>
</table>

With interaction, high vs. low income

<table>
<thead>
<tr>
<th></th>
<th>OLS, yearly</th>
<th>IV, yearly</th>
<th>OLS, permanent</th>
<th>IV, permanent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope, $\beta$</td>
<td>.325 (13.64)**</td>
<td>.335 (7.66)**</td>
<td>.409 (15.84)**</td>
<td>.380 (8.07)**</td>
</tr>
<tr>
<td>Self-employed, $\delta_{SE}$</td>
<td>.065 (1.50)</td>
<td>.057 (1.27)</td>
<td>.082 (1.87)*</td>
<td>.062 (1.39)</td>
</tr>
<tr>
<td>High-inc, SE, $\lambda_{SE}$</td>
<td>.009 (0.14)</td>
<td>.006 (0.09)</td>
<td>-.029 (-0.47)</td>
<td>-.013 (-0.21)</td>
</tr>
<tr>
<td>Salary worker, $\delta_{SW}$</td>
<td>.034 (1.01)</td>
<td>.029 (0.84)</td>
<td>.058 (1.72)*</td>
<td>.045 (1.31)</td>
</tr>
<tr>
<td>High-inc, SW, $\lambda_{SW}$</td>
<td>-.085 (-2.07)**</td>
<td>-.082 (-1.95)*</td>
<td>-.109 (-2.64)**</td>
<td>-.097 (-2.31)**</td>
</tr>
<tr>
<td>Implied $k_{SE,LI}$</td>
<td>1.22 (1.93)</td>
<td>1.19 (1.50)</td>
<td>1.27 (4.17)**</td>
<td>1.23 (2.67)</td>
</tr>
<tr>
<td>Implied $k_{SE,HI}$</td>
<td>1.26 (2.28)</td>
<td>1.21 (1.50)</td>
<td>1.19 (2.13)</td>
<td>1.19 (1.59)</td>
</tr>
<tr>
<td>Implied $k_{SW,LI}$</td>
<td>1.10 (0.78)</td>
<td>1.08 (0.54)</td>
<td>1.12 (1.67)</td>
<td>1.09 (0.90)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.445</td>
<td>.446</td>
<td>.451</td>
<td>.453</td>
</tr>
<tr>
<td>1st stage F-stat</td>
<td>253.9</td>
<td></td>
<td>324.9</td>
<td></td>
</tr>
<tr>
<td>Sargan ($p$–value)</td>
<td>.279</td>
<td></td>
<td>.595</td>
<td></td>
</tr>
<tr>
<td>Wu-H ($p$–value)</td>
<td>.856</td>
<td></td>
<td>.337</td>
<td></td>
</tr>
</tbody>
</table>

Obs.                      | 4207         | 4055       | 4213          | 4057          |

$p < 0.1$ * $p < 0.05$ ** $p < 0.01$***

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, treating the variances of the correction term as non-stochastic. Regressions include 4 controls: age, age squared, number of adults and children. Additional instruments for IV: size of house, dummy for higher education, and log of capital income.
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. Here we pool the information from the two surveys, which implies that we exploit information from about approximately 3000 respondents, approximately 1000 from the 2006 survey and 2000 from 2009 survey.

To obtain information about the determinants for being involved in non-recorded transactions, we estimate a probit model for purchasing in the informal market, corresponding to Equation 10,

$$\Pr \left( c_h = 1 \mid Z' \right),$$

(20)

where $c_h = 1$ for individuals that report they have purchased in the hidden market, i.e. have $c_h > 0$, and where (as above) $Z'$ symbolizes control variables. In the estimations we restrict the sample to household income below 3.5 million NOK\textsuperscript{16}, leaving us with 3259 observations. Table 2 reports estimates from the probit estimation, showing that income ($x$), gender ($m$) and region ($r$) are significant explanatory characteristics for the probability of participating in the hidden market: The probability of being in the informal market increases with income, males ($m = 1$) have a higher probability than females to purchase in the informal market, and location matters, as people in eastern part of Norway, including the capital Oslo ($r = 1$), has a higher probability for being involved in hidden transactions.

To provide a depiction of how these estimates are used to compute the tax evasion corrected incomes, in Table 3 we show probabilities for three different household income levels, dependent on gender.\textsuperscript{17} For example, the probability of entering the informal market increases from 0.16 to 0.23 when the income of the household goes up from 600 000 NOK to 1.4 million NOK.

5 A tax evasion controlled distribution of income

5.1 Imputation of results\textsuperscript{18}

In order to translate the empirical findings from the previous section into effects on the overall income inequality, we use the tax-benefit model LOTTE

\textsuperscript{16}When using exchange rates for 2009, this corresponds to approximately $560,000.

\textsuperscript{17}Non-significant explanatory variables are not used, and the region parameter is set to its average.

\textsuperscript{18}Note that many of the assumptions that are applied in the following will be subject to extensive sensitivity testing in future versions of the paper.
Table 2: Results from probit estimation, pooled data, 2006 and 2009

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross household income,(^1)</td>
<td>.0003 (3.90)***</td>
</tr>
<tr>
<td>Age</td>
<td>-.002 (.002)</td>
</tr>
<tr>
<td>Male</td>
<td>.147 (2.66)***</td>
</tr>
<tr>
<td>Education</td>
<td>.067 (.060)</td>
</tr>
<tr>
<td>Self-employed</td>
<td>.048 (.122)</td>
</tr>
<tr>
<td>Eastern region dummy</td>
<td>.302 (4.43)***</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.29 (.111)</td>
</tr>
</tbody>
</table>

Likelihood Ratio 47.66
No of obs. 3259

\(^1\) Income measured in 1,000 Norwegian kroner

\(^\) z-statistics based on standard deviation reported in parentheses

---

Table 3: Computed probabilities for three different household income levels

<table>
<thead>
<tr>
<th>Income Level</th>
<th>600 000 NOK</th>
<th>1 000 000 NOK</th>
<th>1 400 000 NOK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Pr (c_h^* = 1 \mid Z')</td>
<td>.161</td>
<td>.127</td>
<td>.192</td>
</tr>
</tbody>
</table>

Income measured in 2009 values
This allows us to trade evasion behavior from both sides of the market into economic gains and to keep track of the balance between amounts at the supply and demand side. Then we are able to discuss how the "hidden-economy-controlled" income distribution compare to the official one: is it less or more equal? In addition we can assess who gains most from evasion.

When imputing the unreported income we use the OLS-estimates for permanent income from Table 1. This means that income for households containing at least one self-employer is adjusted according to, \( k \) values, 1.27 for low income household and 1.19 for those with observed income above median. Correspondingly, an income correction has been allocated to low-income salary workers, conditional on type of industry, based on the estimate of \( k_{SW,LI} \) in Table 1.

To implement the demand side estimates, given that the probabilities in Table 2 have been estimated for individuals, we assume individual actions in the households are independent and compute the household probability from the individual probabilities. As denoted by Equation (13) the sum of unrecorded income must equate the sum of hidden expenditures. We assume that demand for the hidden good can be represented by

\[
\ln c_h = a + \varphi \ln Y^*_h, \tag{21}
\]

such that \( \varphi \) can be interpreted as the income elasticity and corresponds to \( d_{Y^*_h} \) as referred to in Section 3.2. As we do not have the possibility to estimate Equation (21) directly, we proceed by using knowledge about income elasticities and Equation (13) to calibrate \( a \). It is probably a reasonable point of departure to assume \( \varphi \) to be around 1.

The expected purchase for a certain household can then be written as

\[
\Pr\left(c^*_h = 1 \mid Z'\right) \exp(a + \varphi \ln Y^*_h), \tag{22}
\]

and combining this with the equilibrium condition in (13) results in

\[
\sum x_h = \sum \Pr\left(c^*_h = 1 \mid Z'\right) \exp(a + \varphi \ln Y^*_h). \tag{23}
\]

Thus, Equation (23) can further be rewritten to

\[
a = \ln \left( \frac{\sum x_h}{\sum \Pr\left(c^*_h = 1 \mid Z'\right) Y^*_h \varphi} \right) \tag{24}
\]

which give us the possibility to calibrate the parameter \( a \) for any given \( \varphi \). As a point of departure, we assume that the consumer price in the hidden market is 20 percent lower than in the regular market. The consumer price in
the informal market is then equal to the pre-VAT price in the regular (VAT rate is 25%), i.e., we have $p_r = p_h$. Further we fix the income elasticity, $\varphi$, at the value 1.

### 5.2 More equal distribution of income?

In Tables 4 and 5 we have adjusted the measured distribution of income to control for income being underreported and some consumers taking advantage of lower prices in the informal market.\(^{19}\) First, Table 4 presents the distributional effect of the income underreporting of the supply side. We see that, on average, as measured by post-tax income among all households, the income underreporting corresponds to 97.2 percent of average post-tax income. This means that post-tax income would be 29 billion NOK higher if people reported their income correctly.

More importantly, given the perspective of the present study, the overall picture is that the rate of underreporting increases with income, and, in particular, the underreporting rates are highest in deciles 9 and 10. The share of reported income to the adjusted income ranges from 99.6% in the second decile to 94.2 percent in the 10th, a relative increase of about 5.7 percent. 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, we find opposite effects with respect to the overall income distribution. This follows from the placement of the tax evaders in the overall income distribution. There are other groups (than self-employed and specific groups of wage earners) dominating in the low end of the income distribution. As seen in the last row, the effect is confirmed by the Gini coefficient.

Next we add the consumer benefits to the income distribution. As seen in Table 5, and as expected given the results of Table 2, this adds to a more unequal "true" income distribution. The share of reported income to the adjusted ranges from 93.5 percent to 99.1 percent, and the Gini coefficient raises further. Because we assume $\varphi = 1$ in this case, this result is driven only by the fact that the probability of purchasing hidden services increases with income.\(^{20}\)

Another way to illustrate the effect of tax evasion on the income distribution is to show how the economic gains are allocated between the two sides of

\(^{19}\)We rank the households according to their "true" income. If we used the reported income concept, those evading would be placed lower in the distribution than they actually are, discussed in Johns and Slemrod (2010)

\(^{20}\)Future sensitivity tests will examine the implications of empirical assumptions.
Table 4: Income distribution of 2009 controlled for underreporting

<table>
<thead>
<tr>
<th>Deciles</th>
<th>Reported after-tax income</th>
<th>Adjusted for underreporting</th>
<th>Reported after-tax income as share of adjusted (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>93112</td>
<td>93688</td>
<td>99.39</td>
</tr>
<tr>
<td>2</td>
<td>172946</td>
<td>173680</td>
<td>99.58</td>
</tr>
<tr>
<td>3</td>
<td>223609</td>
<td>225153</td>
<td>99.31</td>
</tr>
<tr>
<td>4</td>
<td>283025</td>
<td>285232</td>
<td>99.23</td>
</tr>
<tr>
<td>5</td>
<td>343196</td>
<td>346847</td>
<td>98.95</td>
</tr>
<tr>
<td>6</td>
<td>415406</td>
<td>424715</td>
<td>97.81</td>
</tr>
<tr>
<td>7</td>
<td>505754</td>
<td>514987</td>
<td>98.21</td>
</tr>
<tr>
<td>8</td>
<td>598407</td>
<td>610730</td>
<td>97.78</td>
</tr>
<tr>
<td>9</td>
<td>710148</td>
<td>731984</td>
<td>97.02</td>
</tr>
<tr>
<td>10</td>
<td>1069290</td>
<td>1135595</td>
<td>94.16</td>
</tr>
<tr>
<td>All</td>
<td>441489</td>
<td>454261</td>
<td>97.19</td>
</tr>
</tbody>
</table>

Gini 0.356 0.364

Households sorted by adjusted after-tax income in 2009. Values refer to mean in deciles

Table 5: Income distribution of 2009 controlled for underreported income and consumer gains

<table>
<thead>
<tr>
<th>Deciles</th>
<th>Reported after-tax income</th>
<th>Adjusted for underrep. and consumer gains</th>
<th>Reported after-tax income as share of adjusted (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>93112</td>
<td>93961</td>
<td>99.10</td>
</tr>
<tr>
<td>2</td>
<td>172946</td>
<td>174189</td>
<td>99.29</td>
</tr>
<tr>
<td>3</td>
<td>223609</td>
<td>225893</td>
<td>98.99</td>
</tr>
<tr>
<td>4</td>
<td>283025</td>
<td>286331</td>
<td>98.85</td>
</tr>
<tr>
<td>5</td>
<td>343196</td>
<td>348420</td>
<td>98.50</td>
</tr>
<tr>
<td>6</td>
<td>415406</td>
<td>427011</td>
<td>97.28</td>
</tr>
<tr>
<td>7</td>
<td>505754</td>
<td>518216</td>
<td>97.60</td>
</tr>
<tr>
<td>8</td>
<td>598407</td>
<td>614989</td>
<td>97.30</td>
</tr>
<tr>
<td>9</td>
<td>710148</td>
<td>737620</td>
<td>96.28</td>
</tr>
<tr>
<td>10</td>
<td>1069290</td>
<td>1147910</td>
<td>93.15</td>
</tr>
<tr>
<td>All</td>
<td>441489</td>
<td>457454</td>
<td>96.51</td>
</tr>
</tbody>
</table>

Gini 0.356 0.365

Households sorted by adjusted after-tax income in 2009. All values refer to decile mean.
Consumer gains corresponds to VAT (price reduction assumed 20%)
<table>
<thead>
<tr>
<th>Deciles</th>
<th>Supplier gain</th>
<th>As share of adjusted income (%)</th>
<th>Consumer gain</th>
<th>As share of adjusted income (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>217</td>
<td>0.23</td>
<td>273</td>
<td>0.29</td>
</tr>
<tr>
<td>2</td>
<td>437</td>
<td>0.25</td>
<td>509</td>
<td>0.29</td>
</tr>
<tr>
<td>3</td>
<td>1124</td>
<td>0.50</td>
<td>740</td>
<td>0.33</td>
</tr>
<tr>
<td>4</td>
<td>2019</td>
<td>0.71</td>
<td>1099</td>
<td>0.38</td>
</tr>
<tr>
<td>5</td>
<td>2966</td>
<td>0.85</td>
<td>1573</td>
<td>0.45</td>
</tr>
<tr>
<td>6</td>
<td>3676</td>
<td>0.86</td>
<td>2296</td>
<td>0.54</td>
</tr>
<tr>
<td>7</td>
<td>3615</td>
<td>0.70</td>
<td>3229</td>
<td>0.62</td>
</tr>
<tr>
<td>8</td>
<td>4864</td>
<td>0.79</td>
<td>4259</td>
<td>0.69</td>
</tr>
<tr>
<td>9</td>
<td>8850</td>
<td>1.20</td>
<td>5636</td>
<td>0.76</td>
</tr>
<tr>
<td>10</td>
<td>27904</td>
<td>2.43</td>
<td>12315</td>
<td>1.07</td>
</tr>
<tr>
<td>All</td>
<td>5567</td>
<td>1.22</td>
<td>3193</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Households sorted by adjusted after-tax income in 2009.
Supplier gain equals saved income taxes
Consumer gain equals VAT
Table 7: Tax revenues and compliance rate for direct taxes, 2009

<table>
<thead>
<tr>
<th>Deciles</th>
<th>Actual tax revenue</th>
<th>Full compliance</th>
<th>Compliance rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9499</td>
<td>9716</td>
<td>97.77</td>
</tr>
<tr>
<td>2</td>
<td>22170</td>
<td>22607</td>
<td>98.07</td>
</tr>
<tr>
<td>3</td>
<td>49716</td>
<td>50840</td>
<td>97.79</td>
</tr>
<tr>
<td>4</td>
<td>72061</td>
<td>74080</td>
<td>97.27</td>
</tr>
<tr>
<td>5</td>
<td>90427</td>
<td>93393</td>
<td>96.82</td>
</tr>
<tr>
<td>6</td>
<td>117449</td>
<td>121125</td>
<td>96.97</td>
</tr>
<tr>
<td>7</td>
<td>154411</td>
<td>158026</td>
<td>97.71</td>
</tr>
<tr>
<td>8</td>
<td>196530</td>
<td>201394</td>
<td>97.58</td>
</tr>
<tr>
<td>9</td>
<td>258355</td>
<td>267205</td>
<td>96.69</td>
</tr>
<tr>
<td>10</td>
<td>489660</td>
<td>517570</td>
<td>94.61</td>
</tr>
<tr>
<td>All</td>
<td>146028</td>
<td>151595</td>
<td>96.33</td>
</tr>
</tbody>
</table>

Households sorted by adjusted after-tax income in 2009. All values refer to decile mean.

the market for different income levels. This is done in Table 6. For the first deciles, the economic gains are rather modest both in absolute and relative sense, because there are few households with the opportunity of tax evasion in these groups, making the average gains small.

Table 6 also depicts how the effects on the supply and demand side differs with income. For example, at the top of the income distribution, gains from both sides are large, but underreporting of income plays a more important role. In general, the distribution of gains for the consumer side is more equal than the gains at the supply side.

We should note that this, of course, will depend on the assumed splitting of gains. A large price reduction means large benefits to those purchasing and low benefits for those underreporting. As discussed in Section 3.3, this result depends on the shape of the demand and supply curves. If the obstacles for entering the hidden market for the suppliers are low, the prices in the informal income can be driven far down and a large reduction in consumer prices can occur. From a distributional point of view, this could seem like a preferable situation.

Finally, we compare current tax revenues to a hypothetical situation in which all incomes are taxed. In Table 7 we compute personal income tax compliance rates for each decile. A decrease in compliance rates with in-
creased income implies that the tax system becomes less progressive from tax evasion. We see that the compliance rate tends to decrease with income, even though the compliance rates are relatively high in the 7th and the 8th deciles. The mean value for the whole population is 96.3 percent. This corresponds to a difference between the actual tax revenue and the full compliance case of around 13 billion NOK. The lost VAT sums to about 7 billion NOK, so income tax and lost VAT due to this behaviour sums to about 20 billion NOK.

6 Summary

Who gains most from tax evasion, the rich or the poor? In the present paper we have addressed how to account for the fact that much tax evasion behavior involves the participation of more than one taxpayer, and parameterize the distributional consequences using data from Norway. Our preliminary analyses show that the tax-evasion-controlled estimate of income inequality in Norway exhibits more income dispersion than official estimates and that both effects at the supply and demand sides of the market contribute to this.

References


Opinion (2009). Privates kjøp av svart arbeid, Skattedirektoratet. (in Norwegian)


TNS Gallup (2006). Kjøp av svart arbeid i privatmarkedet, Skattedirektorat. (in Norwegian)


