How to Use One Instrument to Identify Two Elasticities

Evelina Gavrilova, Floris Zoutman, Arnt-Ove Hopland, Jarle Møen,
NHH Norwegian School of Economics and
Norwegian Center for Taxation

Preliminary and Incomplete
August 30, 2016

Abstract
By applying economic theory we are able to use one instrument to identify structurally the responses in both supply and demand. We use an implicit assumption taken in tax models since Ramsey (1927) and show that it can serve as an additional exclusion restriction for the tax rate as an instrument. We name it the Ramsey Exclusion Restriction. We present a case study where we use quasi-experimental variation in changes of the payroll tax rates for manufacturing sectors within Norway, to estimate the incidence of the payroll tax. Our data is a panel of the universe of manufacturing firms in Norway between 1996-2012. We find that a 1 percent increase in the payroll tax increases the decreases the gross wage cost by 0.4-0.6 percent indicating that employees pay for 40-60 percent of the payroll tax on average. We further use the changes in the payroll tax in combination with our incidence equation to instrument for changes in the total wage cost per hour, and find labor demand elasticities in the range between -3.5 and -6.

1 Introduction
It is well known that a single instrument can generally not be used to estimate both the demand and supply elasticity.\footnote{See e.g. Koopmans (1949) for a proof} Intuitively, tracing out the slope of the demand curve requires exogenous variation in the supply curve, and vice versa. However, in specific cases economic theory enables us to restrict the structural demand and supply equations, such that two elasticities are identified with a single instrument. In this paper we show that a relatively simple insight from taxation theory allows us to plausibly restrict supply-demand systems when the instrument is a (monotonous transformation of) a tax rate levied on the good. Hence, exogenous variation in the tax rate allows identification of both the demand and the supply elasticity.

Our result is best understood in the context of an ad-valorem tax on a good. Figure 1 displays the effect of an increase in such a tax on the price and traded quantity. In the left panel the horizontal axis represents the traded quantity and the vertical axis represents the price excluding the tax. As can be seen, in this coordinate system the tax increase shifts the demand curve inward from $D_0$ to $D_1$. For any given before-tax price, the tax increase results in a higher price for consumers, and hence reduces demand. The tax reform has no independent effect on supply, and hence, satisfies the standard exclusion restriction. Therefore, the tax rate is a valid instrument for estimating the supply elasticity.

\*We would like to thank Etienne Lehmann, Andreas Peichl, Dirk Schindler and Guttorm Schjelderup for very helpful suggestions and comments. Further, we would like to thank participants of the IIPF conference in Dublin. We acknowledge a contribution by the Research Council of Norway Grant Number 239120.
The right panel shows the exact same tax reform, but the vertical axis now portrays the price including the tax, instead of the price excluding the tax. Fixing the price after taxation implies that demand is unaffected, while supply moves inward. Intuitively, when the tax rate increases it becomes more expensive to produce the same quantity for a given after-tax price. Hence, a simple transformation of variables reveals that the exact same tax reform also allows identification of the demand elasticity.

The result in figure 1 follows from an assumption that is (implicitly) taken in virtually every economic model of taxation since Ramsey (1927). Ramsey (1927) asserts that a tax on a good only affects demand through its impact on the after-tax price of the good. Hence, a 1-percentage increase in the before-tax price and a 1-percentage increase in the net-of-tax rate result in the exact same reduction in demand (in the context of an ad-valorum tax the net-of-tax rate is defined as 1+ the tax rate). This assumption is consistent with rational behavior, as an increase in the before-tax price and the net-of-tax rate have the same effect on the budget constraint of the consumer. Moreover, in many real-life situations the consumer only observes the price after taxation.

In an econometric setting Ramsey’s assumption functions as an additional exclusion restriction. Under this additional exclusion the net-of-tax rate is a valid instrument for estimating both the demand and the supply elasticity. In honor of Frank Ramsey we name the additional exclusion restriction the Ramsey Exclusion Restriction, and abbreviate it by the RER throughout the text.

In addition to instrument validity, identification also requires that the net-of-tax rate is a relevant instrument. Instrument relevance can be subdivided into two components. First, a change in the net-of-tax rate should correlate to the price prior to taxation. If the instrument does not, the burden of the tax is entirely borne by the demand side. Hence, the net-of-tax rate does not provide variation that allows identification of the supply elasticity. Second, the net-of-tax rate should correlate to the price including the tax rate. If it does not, the tax is entirely borne by the producers, and can hence not be used to estimate the demand elasticities. Both elasticities are always (asymptotically) identified if the incidence of the tax is shared between the demand and the supply side.

We formally show that our result scales up to a setting where multiple goods are traded, and each good faces an ad-valorum tax levied on the demand side. If the net-of-tax rates on each good satisfy both

---

2A very similar set of figures is often used in textbooks to explain the difference between a tax levied on the demand and the supply side of the market. See for instance figure 19.7 in Gruber (2010). To clarify, both panels in figure 1 show a tax that is levied on the demand side. The only difference is the coordinate system in which the tax reform is represented.

3The preceding paragraph describes the strong version of RER. The weak version of RER states that there must exist a monotonic transformation of the tax rate such that a 1-percent change in the monotonic transformation of the tax rate has the same effect on demand as a 1-percent change in the price. In practice, the net-of-tax rate is the most likely possible monotonic transformations.
the standard exclusion restriction and the RER the net-of-tax rates are valid instruments for the entire set of demand - and supply (cross)-elasticities. If they in addition provide independent variation in the vector of before- and after-tax prices, the instruments are also relevant.

We provide a simple estimation method that allows one to estimate the set of supply and demand (cross) elasticities. We divide the estimation of the elasticities into two different two-stage regressions. Analogue to the left panel of figure 1, the first stage of the first regression uses the vector of prices excluding the tax as the endogenous variable, and the net-of-tax rates on each good as the instruments. The dependent variables in the second stage equation are the equilibrium quantities of each of the goods. In the second stage the coefficients on the vector of prices measure the supply (cross) elasticities. To estimate the demand elasticities we do exactly the same but replace the vector of prices excluding the tax with the vector of prices including the tax. The coefficients on the prices including the tax denote the demand elasticities. Both regressions can be estimated using standard 2SLS or 3SLS methods. In this setting regular strength-of-instrument tests can be applied to test whether the instruments are indeed relevant, and to assess their strength in finite samples.

Further we provide two extensions to our main result. First, if the tax is levied on the supply instead of the demand side, a simple transformation of the prices leads to the same system of equations. Hence, our result does not depend on which party bears the statutory burden of the tax. Second, our result generalizes to both specific and non-linear taxes. With a specific, rather than an ad-valorum tax, variation in the tax rate allows one to identify the slopes of the (linear) supply and demand curve using only one instrument. If the tax is non-linear it cannot serve as an instrument, because the non-linear tax rate depends on the traded quantity, which results in reverse causality. However, the literature on the on the elasticity of taxable income has created IV approaches to deal with this issue (see e.g. Gruber and Saez (2002), Kopczuk (2005), Weber (2014)), and these can readily be combined with our approach. Hence, our approach can also identify the demand and supply elasticities in case the tax rate is non-linear.

Our result has important implications for welfare analysis. In two seminal articles Harberger (1964b,a) shows that the deadweight loss triangle can be calculated through the information obtained from a reduced-form regression between the net-of-tax rate and the traded quantity. In response a large literature That is, the reduced-form elasticity between the net-of-tax rate and the traded quantity is a sufficient statistic for welfare analysis. Harberger’s result has motivated an extensive empirical literature that aims to estimate the reduced-form elasticity between the net-of-tax rate, and equilibrium quantities or prices. We contribute by showing that the reduced-form elasticities cannot be a sufficient statistic for welfare analysis unless the RER holds. Hence, reduced-form welfare analysis is valid under the same set of conditions that allow one to estimate a structural demand-and-supply model with a single instrument.

The structural analysis has several advantages over reduced-form analysis. First, we show that the RER is testable in a structural framework if additional instruments are available. Contrary, the RER cannot be tested in reduced-form analysis. Second, we show that reduced-form elasticities may depend on market circumstances such as price rigidities, as well as on identification strategies. Our structural approach produces consistent and comparable elasticity estimates across different market circumstances, and identification strategies.

---

Conceptually, this first two-stage regression is similar to the literature that estimate (labor) supply elasticities through exogenous variation in the tax rates. (e.g. Burtless and Hausman (1978); Aaberge et al. (1995); Hausman and Newey (1995); Blundell et al. (1998); Aaronson and French (2009). For instance, Blundell et al. (1998) uses variation in the tax rate resulting from reforms and non-linearity in the tax rates to instrument for the after-tax wage rate. However, we expand upon this methodology by showing that the same variation can also be used to estimate demand elasticities.

Feldstein (1999); Chetty (2009b,a); Hendren (2013); Doerrenberg et al. (forthcoming) discuss the conditions under which the reduced-form elasticity is indeed a sufficient statistic for welfare analysis.

A by no means exhaustive list includes Hamermesh (1979); Gruber (1994); Saez et al. (2012); Lehmann et al. (2013) who estimate the elasticity of employment and wages with respect to the payroll tax. Blomquist (1983); Eissa (1995) estimate the effect of income tax reforms on labor supply. Feldstein (1995); Gruber and Saez (2002); Saez (2010); Lehmann et al. (2013); Kleven and Schultz (2014) estimate the elasticity of taxable income with respect to the income tax. Best and Kleven (2013) who estimate the effect of stamp taxes on housing prices and transactions and Gruber and Köszegi (2001) estimate the effect of excise taxes on cigarette consumption.
We apply our method to a labor-market setting in Norway. Since 1996 there have been several reforms in the payroll tax. We exploit their quasi experimental variation and estimate the incidence of the payroll tax. We use a feature of the data, that allow us to estimate the wage rate per hour and use it as a dependent variable in our estimations. We find that the burden of the tax is shared between employers and employees, contrary to conventional wisdom, which dictates that the incidence should fall fully on employees.

Our application contributes to the empirical literature on payroll taxation. Our results on incidence strongly contrast the recent results by Saez et al. (2012), Lehmann et al. (2013) as both papers show evidence that the entire incidence of the payroll tax remains with the employer. Using a quasi-experimental reform in Sweden, Bennmarker et al. (2009) finds similarly that the payroll tax cut is mostly paid for by employers. However, by evaluating a Finnish reform Korkeamäki and Uusitalo (2009) find that the majority of the incidence resides with the employee respectively. This finding is echoed by earlier studies in the US (e.g. Gruber 1994, Anderson and Meyer 1997, Murphy 2007) which show that several taxes and subsidies levied at the employer level are mostly, but not entirely, shifted onto employees. Gruber (1997) studies a payroll tax reform in Chile and finds that there the payroll tax is entirely shifted onto employees. However, this may be the result of a simultaneous reform in the income tax that was meant to offset the change in revenue induced by the payroll tax reform. Our results are in line with earlier on the incidence of payroll taxation. In Norway Carlsen and Johansen (2005), Dyrstad (1992), Dyrstad and Johansen (2000) and Johansen (2002) include the payroll tax rate as a control variable when explaining the wage cost per workers. Their results indicate the burden of the payroll tax is split about equally between employer and employee.

In each of these studies the dependent variable used is the wage per worker, instead of the wage rate per hour. Therefore, the results may be biased by behavioral responses of the employer and/or the employee with respect to the payroll tax. As far as we know, the only paper in the literature that does use the wage rate per hour as the dependent variable is Johansen and Klette (1997). However, their study was never concluded. Their preliminary results are in line with our preliminary results.

Several previous studies have considered the role of payroll taxes in determining labor demand and employment. Korkeamäki and Uusitalo (2009) and Bohm and Lind (1993) find no effects on employment following payroll tax reforms, consistent with the incidence being borne mostly by the employee. In contrast, Bennmarker et al. (2009) initially find no employment effects, when evaluating the incidence. However, once they account for firm numbers, they find positive effects on firm entry and generally positive effect on employment, while the incidence becomes fully borne by the employer. We are not aware of studies that have evaluated other margins of behavioral responses of firms.

2 Methodology

2.1 Estimating the demand and supply elasticity for a single good using a single instrument

We first turn to an instructive example in a market for a single good. Assume you have panel data on equilibrium quantity of good $Y_{it}$, as well as its price excluding the tax $P_{it}$. Here, the cross-sectional indicator $i$ may, for instance, indicate specific regions, firms or individuals. Assume the good faces an ad-valorem tax rate $\tau_{it}$ and assume the tax is levied on the demand side. Assume there is exogenous variation in the tax rate, possibly after controlling for a $K$-vector of control variables $x_{it}$. In this section we take the exogenous variation in the tax as a given. In the real world the most likely source of variation is a natural experiment where exogenous variation in the tax can be isolated through either a difference-in-difference, or a regression discontinuity design. $x_{it}$ must at the very least contain a constant, but may additionally contain time and/or cross-sectional fixed effects. It may also contain controls for exogenous income that

---

7See section 3 in Bennmarker et al. (2009) for a more detailed overview of all papers in the literature.
allow simultaneous estimation of income effects, as well as other variables that may correlate to both the tax rate and the endogenous variables.

Because the tax is levied on the demand side it only affects supply through its influence on the price $P_{it}$. As such, the instrument shifts the demand curve along the supply curve, but does not appear directly in the supply equation. Initially, we make no restrictions on how the instrument affects the demand curve. As such, if we assume demand and supply are assumed are log-linear in prices they can be represented by the following equations:

\begin{align*}
y_{it} = \varepsilon^D p_{it} + \gamma z_{it} + \Gamma^D x_{it} + \nu^D_{it}, \tag{1} \\
y_{it}^S = \varepsilon^S p_{it} + \Gamma^S x_{it} + \nu^S_{it}. \tag{2}
\end{align*}

where $y_{it}$ and $p_{it}$ denote the log of $Y_{it}$ and $P_{it}$. $\varepsilon^D$ ($\varepsilon^S$) denotes the demand (supply) elasticity. Standard economic theory implies that $-\varepsilon^D, \varepsilon^S > 0$, but our result does not rely on this assumption. $z_{it} \equiv f(\tau_{it})$ is a monotonous transformation of the ad-valorum tax rate chosen ensuring that $Y_{it}$ becomes log-linear in $f(\tau_{it})$. The role of the function $f(\cdot)$ will become more clear below. $\gamma$ is the coefficient on the instrument $z_{it}$. $\Gamma^D (\Gamma^S)$ is the vector of coefficients belonging to the control variables in the demand (supply) equation. Finally, $\nu^D_{it} (\nu^S_{it})$ denotes the disturbance term in the demand (supply) equation.

To see that we cannot identify both the demand and supply elasticity without making additional assumptions, consider the reduced-form equations:

\begin{align*}
\begin{bmatrix} y_{it} \\ p_{it} \end{bmatrix} = \pi z_{it} + \Phi x_{it} + \xi_{it}. \tag{3}
\end{align*}

where $\pi$ is a $2 \times 1$ vector of reduced-form coefficients measuring the effect of the instrument on the quantity and price respectively, and $\Phi$ is a $2 \times K$ matrix of reduced-form coefficients for each of the control variables. Since $z_{it}$ and $x_{it}$ are jointly exogenous, the coefficients $\pi$ and $\Phi$ can be recovered by estimating (3) using OLS.

The coefficients in the structural system of equations (1-2) are identified if they can be expressed in terms of reduced-form coefficients $\{\pi, \Phi\}$ (see e.g. Hausman (1983) for a general treatise on identification in structural equation models). Substituting the system of equations (1-2) into the reduced-form expression we find the following link between the reduced-form coefficients and structural coefficients of interest:

\begin{align*}
\pi = \begin{bmatrix} \varepsilon^S \gamma \\ \frac{\varepsilon^S - \varepsilon^D}{\varepsilon^S - \varepsilon^D} \end{bmatrix} \cdot \gamma
\end{align*}

The right-hand side of the expression is composed of 3 structural coefficients, $\gamma$, $\varepsilon^S$ and $\varepsilon^D$. However, the vector $\pi$ on the left-hand side only has two elements. It is therefore impossible to identify all three structural parameters. One can only solve for the supply elasticity by noting that:

\begin{align*}
\varepsilon^S = \frac{\pi_{zy}}{\pi_{zp}}, \tag{5}
\end{align*}

where $\pi_{zy}$ denotes the reduced-form coefficient that links the instrument to the traded quantity, and $\pi_{zp}$ denotes the coefficient between the instrument and the price. Hence, as long as $\pi_{zp} \neq 0$, the supply elasticity is identified. However, there is no way to back out the demand elasticity.

The above reveals the standard argument for why we need at least two instruments to estimate both the demand and supply elasticity. However, using economic theory we can further restrict the system of equations (1-2). If consumers act rationally the tax only affects demand through its impact on the after-tax price, $P^\tau_{it} \equiv (1 + \tau_{it})P_{it}$, but does not additionally have an independent effect on behavior. Even if consumers do not act rationally, this assumption is still likely to hold in settings where consumers only observe the price after taxation. The assumption that a tax affects behavior through the price, and has no further additional impact on behavior is applied in virtually all economic models of taxation dating
back to Ramsey (1927).\footnote{See for instance Harberger (1964a,b); Mirrlees (1971); Diamond and Mirrlees (1971). To our knowledge the only exceptions are the case where the tax on the good is not full salient, e.g. Chetty et al. (2009) and the case where (part of) the tax may be avoided/evaded e.g. Chetty (2009a).} The assumption that the tax rate only affects demand through its effect on the price after taxation is the RER. Under the RER log-linear demand can be simplified as follows:

\[ y_{it} = \varepsilon^D p_{it} + \Gamma^D x_{it} + \nu^D_{it}, \]

(6)

\[ y_{it} = \varepsilon^D p_{it} + \varepsilon^D \log (1 + \tau_{it}) + \Gamma^D x_{it} + \nu^D_{it}, \]

(7)

As can be seen, equation (7) is a special case of (1). First, the monotonous transformation \( f(\cdot) \) is chosen such that the instrument equals the log of the net-of-tax rate, \( z_{it} = \log(1 + \tau_{it}) \). Second, the coefficient on the instrument is identically equal to the demand elasticity. Intuitively, for the consumer it does not matter whether the price before taxation \( P_{it} \) or the net-of-tax rate \( 1 + \tau_{it} \) goes up by 1 percent. Both changes have exactly the same effect on the price after taxation.

It is easy to show that the demand and the supply elasticities in the structural system of equations (2,7) are identified. To see this again relate reduced-form coefficients to the structural coefficients:

\[ \pi = \begin{bmatrix} \varepsilon^S & \varepsilon^D \\ \varepsilon^S & \varepsilon^D \end{bmatrix}. \]

(8)

The right-hand side of equation (4) only contains two structural coefficients. Hence, the supply elasticity can still be expressed in terms of the reduced-form coefficients using equation (5). However, in addition it is now possible to express the demand elasticity by noting that:

\[ \varepsilon^D = \frac{\pi_{zy}}{1 + \pi_{zp}}. \]

(9)

Therefore, under RER, both the demand and supply elasticity are identified as long as \( \pi_{zp} \neq 0 \) and \( \pi_{z2} \neq 1 \).

These latter two conditions have a straightforward economic interpretation. If \( \pi_{zp} = 0 \) variation in the tax rate does not affect the price prior to taxation. In that case the entire incidence of the tax falls on the demand side. As such, the instrument does not provide variation in the price that is relevant for the supply side, and hence, the supply elasticity is not identified. By considering equation (8), we can see that \( \pi_{zp} = 0 \) if \( \varepsilon^D = 0 \) and/or \( \varepsilon^S = \infty \).

If \( \pi_{z2} = -1 \) the price after taxation is independent of the tax rate. The incidence of the tax hence falls completely on the supply side. Therefore, there is no variation in the price consumers pay for the good, and it hence becomes impossible to estimate the demand elasticity. By considering equation (8), we can see that \( \pi_{zp} = 0 \) if \( \varepsilon^D = -\infty \) and/or \( \varepsilon^S = 0 \).

The insight that the structural coefficients can be expressed in terms of reduced-form coefficients is not entirely new. Elias (2015) uses equation (9) and (5) for back-of-the-envelope calculations of the demand and supply elasticities in a labor-market setting. However, our contribution is to link this insight to the literature on simultaneous equation modeling. We will show below that our approach allows one to use standard 2SLS or 3SLS methods to estimate the elasticities. As such, our approach does not just give a central estimate for the elasticities as in Elias (2015). It additionally allows the calculation of standard errors and confidence bounds. Moreover, standard tests for instrument strength apply. Our approach can also be extended to allow for overidentification tests if there are at least two instruments available. Finally, in the next subsection we show that the same insight can be generalized to a setting with multiple goods, allowing one to estimate \( J \) demand and supply elasticities using only \( J \) instruments.

### 2.2 Multiple goods

To generalize the result to a setting with multiple goods, consider a demand-supply system of \( J \) goods. Let \( y^j_{it} \) denotes the log quantity of good \( j \). Assume each good faces an ad-valorum tax \( \tau^j_{it} \), and assume...
each of the tax rates contains exogenous and independent variation. Again we use as an instrument a
known monotonous transformation of the tax rate such that \( z^j_{it} = f^j \left( \tau^j_{it} \right) \). The full model that relates
the row vector of log prices \( p_{it} \) to supply and demand is given by:

\[
\begin{align*}
\begin{bmatrix}
y^j_{it} \\
y^j_{it}
\end{bmatrix} &= \begin{bmatrix} p_{it} \varepsilon^D_{it} + z_{it} \gamma^D_{it} + x_{it} \Gamma^D_{it} + \nu^D_{it} \\
p_{it} \varepsilon^S_{it} + x_{it} \Gamma^S_{it} + \nu^S_{it}
\end{bmatrix} \quad \forall \ j = 1, \ldots, J,
\end{align*}
\]

(10)

where \( \varepsilon^D_j \) and \( \varepsilon^S_j \) are column vectors of demand and supply (cross) elasticities of good \( j \) with respect
to each of the prices. \( \gamma^D_j \) denotes a column vector of coefficients for each of the \( J \) instruments. \( \Gamma^D_j \) and \( \Gamma^S_j \) denote
the column vectors of coefficients on each of the \( K \) control variables. \( \nu^D_{it} \) (\( \nu^S_{it} \)) denotes
the disturbance term in demand (supply) equation \( j \). Equations \( \{10,11\} \) represent a supply and demand system in its most general form. Demand and supply of each good can potentially depend upon the entire vector of prices \( p_{it} \). Second, as in the example above we initially do not restrict the coefficients \( \gamma^D_j \).

To link the system of equations \( \{10,11\} \) to the literature on identification in simultaneous equation
models it is useful to stack the equations. Let \( N \) denote the total number of observations. \( y^j \) is the
\( N \times 1 \) vector of observations of the log quantity of good \( j \), and \( p^j \) is the corresponding vector of prices.
\( Y = [y^1, \ldots, y^J, p^1, \ldots, p^J] \) denotes the \( N \times 2J \) matrix of endogenous variables. Similarly, let \( z \) denote
the \( N \times J \) matrix of instruments and \( x \) the \( N \times K \) matrix of control variables. \( Z = [z, x] \) denotes the
\( N \times (J + K) \) matrix of exogenous variables. Finally let \( \nu = [\nu^D_1, \ldots, \nu^D_J, \nu^S_1, \ldots, \nu^S_J] \) denote the \( N \times 2J \)
matrix of disturbance terms. The ordering implies that the demand equations are stacked on the left side,
while supply equations appear on the right side. We can now represent the system of equations, \( \{10,11\} \),
as follows:

\[
YB + Z\Gamma = \nu,
\]

(12)

where \( B \) is a \( 2J \times 2J \) matrix of coefficients for the endogenous variables in \( Y \), and \( \Gamma \) is the \((J + K) \times 2J \)
matrix of coefficients for the exogenous variables.

As in the one-good model, it is again useful to also define the reduced-form equation, which can be
written as:

\[
\begin{align*}
Y + Z\Gamma B^{-1} &= \nu B^{-1}, \\
Y &= Z\Pi + \xi,
\end{align*}
\]

(13)

where \( \Pi = \Gamma B^{-1} \) is a \((J + K) \times 2J \) matrix of reduced form coefficients. We can estimate the coefficients
in \( \Pi \) using OLS, since the variables in \( Z \) are assumed to be jointly exogenous. \( \Pi \) can be partitioned as follows:

\[
\Pi = \begin{bmatrix} \Pi_{zy} & \Pi_{zp} \\ \Pi_{xy} & \Pi_{xp} \end{bmatrix},
\]

where \( \Pi_{zy} \) denotes the partition that contains coefficients between the instruments \( z \) and the goods \( y \), etc.

The structure of \( \{10,11\} \) allows us to restrict the structural coefficients in \( B \) and \( \Gamma \) in several ways.
First, equation \( j \) contains \( y^j \), but does not depend on \( y^k \) for \( k \neq j \). Second, the coefficient on \( y^j \) is \( -1 \).
Hence, we can write:

\[
B = \begin{bmatrix} -I_J & -I_J \\ \varepsilon^D & \varepsilon^S \end{bmatrix},
\]

Second, the instruments do not appear in the supply equations. Hence we can write:

\[
\Gamma = \begin{bmatrix} \gamma^D & 0_J \times J \\ \Gamma^D & \Gamma^S \end{bmatrix}.
\]

It is well known that in this system of equations the matrix of supply elasticities is identified provided
\( \Pi_{zp} \) has full rank (see e.g. Hausman (1983)). However, without further restrictions the demand elasticities
are not identified. The restriction we propose is RER, and is specified formally below. We provide two
definitions of RER. Strong RER restricts $\gamma^D_j$s, and additionally restricts the monotonic transformation to be equal to the log of the net-of-tax rate. This version of RER is the logical multi-dimensional equivalent of the one discussed in section 2.

**Strong Ramsey Exclusion Restriction.** In equation (10) an increase in the instrument $z^j_{it}$ and an increase in the pre-tax price $p^j_{it}$ have the same effect on demand such that $\gamma^D_j = \varepsilon^D_j$ for all $j = 1, \ldots, J$. Moreover, the instruments are defined as:

$$z^j_{it} \equiv \log \left(1 + \tau^j_{it}\right) \quad \forall \quad j = 1, \ldots, J.$$  

We also formulate a slightly weaker version of RER where we restrict the $\gamma^D_j$s, but allow the instrument to be a general monotonic transformation $f(\cdot)$ of the tax rates.

**Weak Ramsey Exclusion Restriction.** The vector of coefficients on the instrument $z_{it}$ is equal to the vector of coefficients on the demand elasticities, $\gamma^D_j = \varepsilon^D_j$, and the instrument is a given monotonic transformation of the tax rate, $z^j_{it} \equiv f^j(\tau_{it})$ for all $j = 1, \ldots, J$.

The difference between the strong and the weak version of RER is that the strong version restricts the monotonic transformation to the net-of-tax rate. Hence, in the strong version of RER a 1 percentage change in the net-of-tax rate yields the same change in demand as a 1 percentage change in the price prior to taxation. With weak RER we can also allow for, for instance, an extra multiplication with a constant $c^j > 0$ such that $z^j_{it} = c^j \log \left(1 + \tau^j_{it}\right)$. The implication of the latter transformation is that a 1-percent change in the before-tax price elicits the same response in demand as a $c$-percent change in the net-of-tax rate. We do however require that the functions $f^j(\cdot)$ are specified up front.

Proposition 1 formally shows that the instruments $z_{it}$ are valid instruments for estimating the full set of demand and supply cross-elasticities. Further, it derives the rank conditions which ensure that $z_{it}$ is also a relevant instrument.

**Proposition 1.** The supply and demand (cross) elasticities in the system of simultaneous equations (10,11) are identified if weak RER applies and $\Pi_{zp}$ and $\Pi_{zp} + I_J$ have full rank.

**Proof.** We prove that all coefficients in the system of equations (10,11) are identified by showing that the structural coefficients in each of the $j = 1, \ldots, 2J$ individual equations are identified. Denote by $\{B_j, \Gamma_j\}$ the $j$-th column of $B$ and $\Gamma$. The two column vectors contain the full set of structural coefficients in equation $j$. Denote the restrictions on the coefficients in matrix form as follows:

$$\begin{bmatrix} \Phi^{B_j} & \Phi^{\Gamma_j} \end{bmatrix} \begin{bmatrix} B_j \\ \Gamma_j \end{bmatrix} = \phi_j,$$

where $\Phi^{B_j}$ ($\Phi^{\Gamma_j}$) is a $g \times 2J$ ($g \times (J + K)$) matrix with restrictions on the coefficients in $B_j$ ($\Gamma_j$), $\phi_j$ is a $g$-vector and $g$ is the total number of restrictions on $B_j$ and $\Gamma_j$. In that case, the structural coefficients in equation $j$ are identified if the system of equations:

$$\begin{bmatrix} \Pi & I_{j+K} \\ \Phi^{B_j} & \Phi^{\Gamma_j} \end{bmatrix} \begin{bmatrix} B_j \\ \Gamma_j \end{bmatrix} = \begin{bmatrix} 0 \\ \phi_j \end{bmatrix},$$

has a single solution, or is overidentified (see Hausman (1983)). The necessary and sufficient condition for this to be true is known as the rank condition and can be written as:

$$\text{rank} \begin{bmatrix} \Pi & I_{j+K} \\ \Phi^{B_j} & \Phi^{\Gamma_j} \end{bmatrix} = 3J + K.$$
To see that the rank condition is satisfied for the supply equations, note again that only one good appears in each equation, and the coefficient on the good is restricted to equal \(-1\). Moreover, instruments do not appear in supply equations. The restriction matrix is thus given by:

\[
[\phi B_j \quad \phi \Gamma_j] = \begin{bmatrix} I_J & 0_{J \times J} & 0_{J \times J} & 0_{J \times K} \\ 0_{J \times J} & 0_{J \times J} & I_J & 0_{J \times K} \end{bmatrix},
\]

for all supply equations. The rank condition for supply equation \(j\) is therefore:

\[
\text{rank} \begin{bmatrix} \Pi_{zy} & \Pi_{zp} & I_J & 0_{J \times K} \\ \Pi_{xy} & \Pi_{xp} & 0_{J \times J} & I_J \\ I_J & 0_{J \times J} & 0_{J \times J} & 0_{J \times K} \\ 0_{J \times J} & 0_{J \times J} & I_J & 0_{J \times K} \end{bmatrix} = 3J + K.
\]

To proof that this condition is satisfied we need to show that the matrix has full rank. Therefore, we consider whether we can use linear row-operations to completely cancel out other rows. If this is not possible the matrix has full row rank, and since it is square, full rank.

Consider the partitions from top to bottom. Rows from the second partition cannot be used to cancel out rows in any of the other partition, as it is the only partition with non-zero elements in the right-most partition. Since the second partition has full row rank we can remove it from consideration. The rank restriction thus simplifies to:

\[
\text{rank} \begin{bmatrix} \Pi_{zy} & \Pi_{zp} & I_J & 0_{J \times K} \\ I_J & 0_{J \times J} & 0_{J \times J} & 0_{J \times K} \\ 0_{J \times J} & 0_{J \times J} & I_J & 0_{J \times K} \end{bmatrix} = 3J.
\]

(14)

The right-most partition no longer contributes to the rank as it consists of zeros, and can therefore be removed from consideration. Furthermore, multiply the second partition from the top in equation (14) with \(-\Pi_{zy}\) and add it to the first partition. We arrive at:

\[
\text{rank} \begin{bmatrix} 0_{J \times J} & \Pi_{zp} & I_J \\ I_J & 0_{J \times J} & 0_{J \times J} \\ 0_{J \times J} & 0_{J \times J} & I_J \end{bmatrix} = 3J.
\]

The second partition from the top has full row rank, and cannot be formed through linear combinations of the other partitions. It can hence be taken out of consideration. The rank condition thus simplifies to:

\[
\text{rank} \begin{bmatrix} \Pi_{zp} & I_J \\ 0_{J \times J} & I_J \end{bmatrix} = 2J,
\]

where we removed the left-most partition as it consists of zeros only. Now multiply the bottom partition with minus one and subtract from the top partition to arrive at:

\[
\text{rank} \begin{bmatrix} \Pi_{zp} & 0_{J \times J} \\ 0_{J \times J} & I_J \end{bmatrix} = 2J.
\]

Both the bottom and the top partition have full row-rank, since we assume \(\Pi_{zp}\) has full rank. Furthermore, we clearly cannot use operations from the first partition to cancel out the second partition or vice versa. Therefore, the rank condition is satisfied.

For demand equations the additional restriction come from RER. The matrix of restrictions on demand equations can be written as:

\[
[\Phi B_j \quad \Phi \Gamma_j] = \begin{bmatrix} I_J & 0_{J \times J} & 0_{J \times J} & 0_{J \times K} \\ 0_{J \times J} & I_J & -I_J & 0_{J \times K} \end{bmatrix},
\]

where we assumed that \(\Phi B_j\) and \(\Phi \Gamma_j\) are square matrices of appropriate dimensions.
The rank condition is hence given by:
\[
\text{rank} \begin{bmatrix}
\Pi_{zy} & \Pi_{zp} & I_J & 0_{J \times K} \\
\Pi_{xy} & \Pi_{xp} & 0_{J \times J} & I_K \\
I_J & 0_{J \times J} & 0_{J \times K} & 0 \\
0_{J \times J} & I_J & -I_J & 0_{J \times K}
\end{bmatrix} = 3J + K.
\]

Applying the same operations as above we can simplify this to:
\[
\text{rank} \begin{bmatrix}
\Pi_{zp} & I_J \\
I_J & -I_J
\end{bmatrix} = 2J,
\]

Finally add the bottom partition to the top partition to arrive at:
\[
\text{rank} \begin{bmatrix}
\Pi_{zp} + I_J & 0_{J \times J} \\
I_J & -I_J
\end{bmatrix} = 2J.
\]

This rank condition is satisfied under the assumption that \(\Pi_{zp} + I_J\) has full rank.

The rank conditions have a useful economic intuition. To see this assume strong RER such that \(z_{it}^j = \log(1 + \tau_{it}^j)\), and consider the reduced-form equations that relate prices to the exogenous instruments:
\[
p_{it} = \log (1 + \tau_{it}) \Pi_{zp} + x_{it} \Pi_{xp} + \xi_{it}^p. \tag{15}
\]

If \(\Pi_{zp}\) does not have full rank variation in the net-of-tax rate does not provide linearly independent variation in the price prior to taxation. Hence, in that case it is impossible to independently estimate the vector of supply elasticities \(\varepsilon^S\).

Now add the vector \(\log (1 + \tau_{it})\) to both sides of the equation to arrive at:
\[
p_{it}^\tau = \log (1 + \tau_{it}) (\Pi_{zp} + I_J) + x_{it} \Pi_{xp} + \xi_{it}^p, \tag{16}
\]

where \(p_{it}^\tau \equiv p_{it} + \log (1 + \tau_{it})\) denotes the price of the good after taxation. Hence, if \(\Pi_{zp} + I_J\) does not have full rank, variation in the instruments does not provide linearly independent variation in the vector of prices after taxation. Therefore, it is impossible to independently estimate the vector of demand elasticities \(\varepsilon^D\). The rank restrictions thus ensure that there is exogenous and independent variation in all prices prior to, and after taxation.

### 2.3 Estimation

In this subsection we show how to estimate the supply and demand elasticities of the system of equations (10) with RER. The estimation strategy is perhaps described by the two panels in figure 1. An exogenous change in the instruments results in variation in before-tax prices that is exogenous to other supply factors. Hence, in order to estimate the supply (cross) elasticities we first run the reduced-form regression described by the right partition of equation (15). This provides us with the reduced-form elasticity between the net-of-tax rate and the before-tax price. We then use instrumented prices to estimate the system of supply equations described by (11), and estimate the supply elasticities in each equation. The system can be estimated using 2SLS where each second-stage equation is estimated independently or through 3SLS where the second-stage equations are estimated jointly.

To estimate the demand elasticities note that analogue to the right panel of figure 1, the instruments provides variation in the price including the tax that is exogenous to other demand factors. Hence, first construct the price including the instrument, \(p_{it}^\tau \equiv p_{it} + z_{it}\). In case of strong RER, \(z_{it} = \log(1 + \tau_{it})\), \(p_{it}^\tau\) is simply the price of the good after taxation = \(p_{it}^\tau\). Substitute \(p_{it}^\tau\) into the demand equations (10):
\[
y_{jit} = p_{it}^{\tau j} \varepsilon^D_j + x_{it} \Gamma_j^D_j + \nu_{it}^{D_j}. \tag{17}
\]
Now to estimate the demand elasticities first estimate the reduced-form expression

\[ p_{zt}^* = \Pi_{zp} z_{it} + \Pi_{xp} x_{it} + \xi_{it}^p \]

(18)
to instrument the prices including the instrument, \( p_{zt}^* \). Second, use instrumented prices to estimate the demand elasticities through equation (17). Again, this system of equation can be estimated using either 2SLS or 3SLS.

As we show in proposition 1 the instruments are relevant for estimating the supply elasticities if the matrix of reduced-form coefficient \( \Pi_{zp} \) has full rank. This matrix is estimated in the first stage equation for the supply elasticities. Hence, standard strength-of-instrument tests (e.g. Sanderson and Windmeijer, 2016) consider the null hypothesis that the first-stage coefficients \( \Pi_{zp} \) do not have full rank against the alternative that \( \Pi_{zp} \) does have full rank, and are therefore appropriate for this application. Moreover, proposition 1 also shows that the instruments are relevant in estimating the demand elasticities provided \( \Pi_{zp} + I \) has full rank. Hence, by a similar argument standard strength of instrument tests can also be used to evaluate the relevance of the instruments for the demand elasticities.

2.4 Extensions

In this subsection we discuss several straightforward extensions to proposition 1. First, consider the case where the tax on the goods is levied on the supply side rather than the demand side. Denote the vector of prices for the demand side by \( \tilde{P}_{it} \). In addition, denote the price the supply side receives for the good by \( P_j^i \equiv \left( 1 - \theta_{it}^j \right) \tilde{P}_{j}^i \), where \( \theta_{it}^j \) denotes the tax rate on good \( j \), and assume the Strong RER Applies such that supply depends on the price of the good after taxation. In that case the supply-demand system is given by:

\[ y_{jt}^i = P_{it}^e D_j + x_{it} \Gamma_{D_j} + \nu_{it}^{D_j}, \]

(19)
\[ y_{jt}^i = P_{it}^e S_j + x_{it} \Gamma_{S_j} + \nu_{it}^{S_j}. \]

(20)

Now define \( \tau_{it} = \frac{1}{1-\theta_{it}} - 1 \) and define the demand equation in terms of \( p_{it} \) rather than \( \tilde{p}_{it} \) to arrive at:

\[ y_{jt}^i = p_{it} \varepsilon_{D_j} + (1 + \tau_{it}) \varepsilon_{D_j} + x_{it} \Gamma_{D_j} + \nu_{it}^{D_j}, \]

which is the original supply-demand system under RER. Hence, the proof of Proposition 1 applies.

Second, assume the government levies a specific rather than an ad-valorum tax on the goods such that the price after taxation is given by: \( P_{it}^\tau = P_{it} + \tau_{it} \). In that case again apply Strong RER, but now specify the supply-demand model in linear, rather than log-linear terms:

\[ Y_{jt}^i = P_{it} \varepsilon_{D_j} + \tau_{it} \varepsilon_{D_j} + x_{it} \Gamma_{D_j} + \nu_{it}^{D_j}, \]

(21)
\[ Y_{jt}^i = P_{it} \varepsilon_{S_j} + x_{it} \Gamma_{S_j} + \nu_{it}^{S_j}. \]

(22)

It can easily be seen that this is a special case of the supply-demand system considered in proposition 1.

Finally, consider the case of a non-linear tax such that the tax rate depends on the traded quantity, \( \tau_{it} (y_{it}) \). Clearly, the tax rate is now endogenous due to reverse causality. However, the literature on the elasticity of taxable income uses IV strategies to overcome endogeneity (see e.g. Gruber and Saez (2002); Kopczuk (2005); Weber (2014)). This IV-approach can readily be combined with our method as we will outline above.

The literature on the elasticity of taxable income studies a tax reform which creates exogenous variation in the tax rates. Due to the non-linearity in the tax rate, the post-reform tax rate itself is endogenous. The idea is to create synthetic instruments by taking the new tax rules and applying them to lagged income
levels. The instruments are plausibly exogenous provided one takes a sufficient number of lags (see Weber (2014) for a discussion).

Hence, to combine this approach with our method first instrument for the tax rate using the first-stage equation outlined in Weber (2014). Second, replace actual with instrumented tax rates in equation (10) and use the estimation method outlined in section 2.3. One concern is that standard errors need to be adjusted for the fact that one uses instrumented rather than actual tax rates. However, this is unlikely to be of any practical concern, since the instruments in the literature on the elasticity of taxable income are typically extremely strong (F-statistics are rarely smaller than 100). Hence, this should not add significant uncertainty to the estimates.

3 Reduced-form vs structural empirical analysis

Most of the recent empirical literature on taxation has focused on estimating the reduced-form elasticities in equation (13) rather than structural demand and supply elasticities. The most important reason for focusing on reduced-form, rather than structural elasticities is stated eloquently by Chetty (2009b):

“The econometric challenge in implementing any of these structural methods is simultaneity: identification of the slope of the supply and demand curves requires 2J instruments.” Yet in this paper we show that the econometric challenge of simultaneity can be overcome when the RER holds, since in that case structural methods can identify the slope of supply and demand curves with only J instruments. Hence, under the RER reduced-form and structural methods require the same number of instruments. This section therefore provides a reappraisal of the advantages of structural analysis vis-a-vis reduced-form analysis is therefore required. This section is divided into 3 parts. Throughout this section we focus on the one-good model introduced in section 3.1. The results in this section can be generalized to a J-good setting but this will complicate the analysis severely, without adding additional insights.

In the first subsection we compare structural and reduced-form welfare analysis. We show that the reduced-form elasticity between the tax rate and the traded quantity is not a sufficient statistic for welfare analysis unless the RER holds. Therefore, reduced-form and structural welfare analysis require the same number of instruments: J if the RER holds, and 2J if it does not. Second, we show that the RER is testable in our structural setting, provided additional instruments are available, whereas it remains fundamentally untestable in a reduced-form framework, independent of the number of instruments. Third, we show the distinct advantage of structural over reduced-form analysis in terms of extrapolation.

3.1 The role of the RER in reduced-form welfare analysis

In two famous articles Harberger (1964b,a) provides the modern foundation of reduced-form welfare analysis. His canonical formula for the marginal increase in the excess burden associated with an ad-valorem tax on good y_{it} is given by:

$$\frac{dEB}{dT_{it}} = \frac{dy_{it}}{dT_{it}} \tau_{it}. \quad (23)$$

In our model, if we assume the instrument is the log of the net-of-tax rate, \(z_{it} = \log(1 + \tau)\), the derivative between the traded quantity and the tax rate is given by reduced-form coefficient \(\pi_{zy}\):

$$\frac{dEB}{dT_{it}} = \pi_{zy} \tau.$$

Hence, the excess burden of a tax can be evaluated by estimating the reduced-form regression (3), and then evaluating the integral:

$$EB = \int_{0}^{\tau} \pi_{zy} \times dx = \frac{\pi_{zy}}{2} \tau^{2}.$$
In other words the reduced-form elasticity, $\pi_{zy}$ is a sufficient statistic for welfare analysis.

As the above-quote by [Chetty 2009b] points out, this apparently provides an important advantage for reduced-form welfare analysis. In the one-good context, the reduced-form elasticity is identified with a single instrument, whereas structural approaches to welfare analysis generally require two instruments.

However, below we show that the reduced-form elasticity cannot be a sufficient statistic for welfare analysis unless the RER holds. To see this, consider excess burden of taxation in the unrestricted supply and demand system given by equations (1,2). The deadweight loss triangle in figure 2 represents the excess burden of the tax when $z_t = z_t^9$. To calculate the size of the excess burden note that, with respect to Laissez-Faire, demand has shifted down by $\gamma \varepsilon S z$. The traded quantity has decreased by $\varepsilon S - \varepsilon D z$. Hence, the area of the deadweight loss triangle is given by:

$$EB = \frac{1}{2} \frac{\varepsilon S \gamma^2 z^2}{(\varepsilon S - \varepsilon D) \varepsilon D}.$$

The marginal excess burden associated with a change in the instrument is therefore:

$$\frac{dEB}{dz} = \frac{\varepsilon S \gamma^2 z}{(\varepsilon S - \varepsilon D) \varepsilon D}.$$

Figure 2: The Deadweight Loss of a tax in an unrestricted supply and demand system

As can be seen, the expression depends on all three structural parameters and therefore cannot be expressed in terms of the reduced-form elasticity. Hence, unless we impose additional restrictions The reduced-form elasticity is not a sufficient statistic for welfare analysis. However, if we impose the strong

\[\gamma \varepsilon S, -\varepsilon D, -\gamma > 0.\]
RER such that $\gamma = \epsilon^D$, and $z = \log(1 + \tau)$ we arrive at Harberger’s canonical formula\textsuperscript{10}

$$\frac{dEB}{d\tau} = \frac{\epsilon^S\epsilon^D}{\epsilon^S - \epsilon^D} = \pi zp \tau.$$ 

Hence, under the strong RER Harberger’s reduced-form analysis applies. However, as we show in section 3 when the RER holds we can also estimate the demand and supply elasticity with a single instrument. The instrument requirement for reduced-form and structural welfare analysis is therefore identical.

### 3.2 Testability of the RER

Under the RER the reduced-form elasticity $\pi zy$ is a sufficient statistic for welfare analysis. However, the RER itself cannot be tested within the context of the reduced-form regression (3). However, if a second instrument is available the RER can be tested in the structural set of equations (1,2). To see this note that testing the weak RER comes down to testing the null hypothesis $\gamma = \epsilon^D$ against the alternative hypothesis $\gamma \neq \epsilon^D$. Denote the second instrument by $z^2_{it}$ and initially assume the instrument shifts the supply curve without affecting the demand curve. In that case (1,2) can be written as:

$$y_{it} = \epsilon^D p_{it} + \gamma z_{it} + \Gamma^D x_{it} + \nu^D_{it},$$
$$y_{it} = \epsilon^S p_{it} + \gamma^2 z^2_{it} + \Gamma^S x_{it} + \nu^S_{it},$$

Without the additional instrument the null hypothesis cannot be tested because the structural coefficients in the demand equation are not identified. However, the additional instrument can identify $\gamma$ and $\epsilon^D$ independently through two-stage least squares. To do this first instrument the price with the $z^2_{it}$ as follows:

$$p_{it} = \pi zp z_{it} + \pi z^2_{it} + \Pi x_{it} + \xi_{it}.$$ 

Second, use instrumented prices, $\hat{p}_{it}$, to estimate the demand equation:

$$y_{it} = \epsilon^D \hat{p}_{it} + \gamma z_{it} + \Gamma^D x_{it} + \nu^D_{it}.$$ 

In this second-stage equation the null hypothesis $\gamma = \epsilon^D$, and hence the RER, can be tested through a standard Wald test.

Now again, assume a second instrument $z^2_{it}$ is available, but now assume the instrument shifts the demand rather than the supply equation. In that case, the demand and supply system can be written as:

$$y_{it} = \epsilon^D p_{it} + \gamma z_{it} + \gamma^2 z^2_{it} + \Gamma^D x_{it} + \nu^D_{it},$$
$$y_{it} = \epsilon^S p_{it} + \Gamma^S x_{it} + \nu^S_{it}.$$ 

Now assume RER holds. In that case, we can rewrite this system of equations using $p^z_{it}$:

$$y_{it} = \epsilon^D p^z_{it} + \gamma^2 z^2_{it} + \Gamma^D x_{it} + \nu^D_{it},$$
$$y_{it} = \epsilon^S p^z_{it} - \epsilon^S z_{it} + \Gamma^S x_{it} + \nu^S_{it}.$$ 

Hence, the RER allows us to reformulate the supply-demand system such that the instrument appears in the supply rather than the demand equation. In this formulation the coefficient on the instrument in the supply equation should equal the supply elasticity. A deviation implies that the RER does not hold. Hence, to test the RER first instrument $p^z_{it}$ using the second instrument:

$$p^z_{it} = \pi zp z_{it} + \pi z^2_{it} z^2_{it} + \Pi^z x_{it} + \xi^z_{it}.$$ 

\textsuperscript{10}In our derivation we used the fact that $dz = d\log(1 + \tau) = d\tau$ on the left-hand side of the equation, and the approximation $\log(1 + \tau) \approx \tau$ on the right-hand side.
Then use instrumented prices to estimate the supply equation:

\[ y_{it} = \varepsilon^S \tilde{p}_{it} + \gamma^S z_{it} + \Gamma^S x_{it} + \nu^S_{it}. \]

In this second-stage equation the RER can be tested through a Wald-test with the null hypothesis \( \varepsilon^S = \gamma^S \).

The fact that the RER can potentially be tested provides an important advantage of structural analysis over reduced-form analysis. There are several reasons why the RER may in practice fail to hold. For instance, Chetty et al. (2009) study the case where the tax on a good is not (fully) salient. They show theoretically, and empirically that in that case a 1-percent change in the before-tax price may not elicit the same change in demand as a 1-percentage change in the net-of-tax rate. In addition, tax evasion or tax avoidance may also result in a failure of the RER (see e.g. Chetty (2009a); Doerrenberg et al. (forthcoming). In this case a change in the price and a change in the net-of-tax rate do not elicit the same change in demand because part of the tax can be evaded or avoided. It is therefore useful to test whether RER holds, and such a test can only be performed in a structural model.

### 3.3 Extrapolation

A final concern related to reduced-form analysis is its inability to make out-of-sample predictions. To see this consider a country that (exogenously) introduces regional differentiation in the ad-valorem tax rate on a specific good \( y_{it} \) by dividing the country into two different tax zones. For simplicity assume the tax rate in region 1 remains constant at 0, whereas the tax rate in zone two increases from 0 to \( \tau \). To estimate the causal effect of the tax rate on the market for good \( y_{it} \) the researcher can either use a regression discontinuity design (RDD) or a difference-in-difference design (DiD). With RDD the researcher would ideally focus in on consumers that live close to, but on different sides on the border. In this context distance from the border forms the running variable. With DiD the researcher divides the two zones into a treatment and a control zones, and uses data on all purchases within the zone, rather than focusing only on consumers close to the border. With the RDD high-tax and low-tax consumers face the same before-tax price because in the region around the border the supply side cannot discriminate its price on the basis of the tax zone in which a consumer resides.\(^{11}\) However, with the DiD analysis consumers within the treated tax zone can shift part of the burden of the tax to the supply side.

Figure XXX represents the variation induced by the tax reform graphically. The left panel shows the effect on the price excluding the tax. The right panel shows the effect on the price including the tax. With RDD the price before taxation is fixed. Hence, the traded quantity for the low-tax group is given by \( y^0 \) whereas the traded quantity for the high-tax group is given by \( y^{RDD} \). For the DiD design the traded quantity of the control group is given by \( y^0 \) whereas the treated group demands \( y^{DiD} \).

As can be seen, the two different identification strategies result in different reduced-form elasticities. If the RER holds the expected value of the traded quantity is \( E[y^{RDD}] = y^0 + \varepsilon^D \tau \). Hence, the central estimate for the reduced-form elasticity between the instrument and the quantity is given by \( \hat{\pi}^{RDD}_{zy} = \frac{E[y^{RDD} - y^0]}{\tau} = \varepsilon^D \). Since the price is unaffected, the central estimate for the reduced-form elasticity between the instrument and the price is zero, \( \pi^{RDD}_{zp} = 0 \). On the other hand, with DiD we arrive at \( \pi^{DiD}_{zy} = \frac{E[y^{DiD} - y^0]}{\tau} = \frac{1}{\varepsilon^S} \varepsilon^D \) and \( \pi^{DiD}_{zp} = \frac{E[p^{DiD} - p^0]}{\varepsilon^S} = \frac{\varepsilon^S}{\varepsilon^D} \). Hence, within the context of the same exogenous variation in the tax rate two different identification strategies result in vastly different reduced-form elasticities.

On the other hand our estimation method would provide a consistent estimate of the demand elasticity across both identification strategies. The supply elasticity is not identified with RDD, because the variation in the tax rate is irrelevant in identifying the supply elasticity. With DiD the estimated supply elasticity is the actual supply elasticity.

\(^{11}\)If the supply side does discriminate the price based on the tax rate consumers face this will typically imply the identifying assumption of the regression discontinuity design does not hold.
4 Application: payroll taxation in Norway

We now turn to an application in the Norwegian labor market. Norway levies a payroll tax (arbeidsavgift) at the employer level. A payroll tax levied at the employer level is conceptually equivalent to an ad-valorum tax levied at the (labor) demand side. Hence, the conceptual framework of section 2 where we discuss an ad-valorum tax in a single-good market carries over to the application. In a labor-market setting effective units of labor hired by a firm act as the quantity variable, and the wage rate per effective unit of labor acts as the price variable. In our data section we discuss how we measure effective units of labor in more detail. Below we first outline the institutional context.

Payroll tax rates in Norway are differentiated across regions. The policy aim of differentiating payroll tax rates by region is to stimulate employment in the Norwegian periphery. Employers in the urban areas in the South of Norway face a tax rate of 14.1 percent. Employers in the rural North of the country are exempt from the tax, and in between are several zones with rates ranging between about 5 and 11 percent. Figure 4 provides an overview of the 7 current tax zones.

Tax revenue from the payroll tax is used to finance social security. One may therefore be concerned that the payroll tax acts as a de facto insurance where higher payroll tax rates lead to higher social security benefits in the future. However, payroll tax rates are regionally differentiated, whereas social security benefits are not. Since we only use this regional variation in our identification strategy the payroll tax can more accurately be viewed as a pure tax, rather than an insurance premium.

Over the years the payroll tax system has been reformed several times. New zones have been introduced, municipalities have shifted between zones, and rates have adjusted. Regional and local governments have no influence on these decisions, because the payroll tax system is fully determined by the central government of Norway. Nevertheless, reforms may be endogenous, because the Norwegian government may, for instance, decide to shift municipalities between zones based on local labor market conditions.

In our application we therefore focus on a rather specific reform that is most likely to be exogenous to local labor market conditions. Before 2000 all employers within a specific zone faced the same tax rate. After 2000 firms in Norway’s most important exporting industries face a 14.1 percent tax rate regardless of their location. The reform was introduced in order to comply with EU legislation, and is hence likely to be exogenous to local labor market conditions.

In 2000 Norway had 5 tax zones. The reform therefore allows us to divide firms in the affected industries into one control and four different treatment groups. The control group consists of firms in the affected industries that already faced a 14.1 percent tax rate because, they are located in urban areas in the South

---

12 During and after the reform period, the payroll tax system faced several other reforms as well. However, these reforms did not affect firms in the exporting industries.
of Norway. Exporting firms in the other 4 tax zones each face a different treatment, depending on the tax rate they faced prior to the reform. Firms in the rural North faced the largest shock, because their payroll tax rate increased from 0 to 14.1 percent. The shock in the rural areas in the South was a lot smaller at 3.5 percentage points.

Part of the 2000 reform was undone in 2007, as some of the industries affected by the first reform were again allowed to pay the tax rate corresponding to their location. By that time the number of tax zones had also increased from 5 to 7. As a result, the 2007 reform divides the country into one control, and six treatment groups. Figure ?? gives an overview of the average tax rate industries affected by the 2000 reform pay over time. Table XX provides a full list of the firms affected by the payroll tax reform.

4.1 Labor Markets in Norway

4.2 Data description

We use firm plant data from the Annual Manufacturing Census of Statistics Norway, covering the period from 1996 to 2012. The dataset contains administrative data on plants, defined as production facilities for a given firm. For the purposes of our analysis we limit our sample to plants with multiple employees in industries that have been directly affected by the payroll tax reform. The treatment group consists of plants in payroll zones 2, 3, 4 and 5 which were affected by the reform, while the control group consists of plants in zone 1.

Table [1] presents summary statistics for treatment and control groups of plants. A treated plant paid an yearly wage rate of 194 Norwegian Krones. It had 40 employees on average, less than the plants in the control group, and it paid an yearly wage of NOK 312 thousand on average. For both treatment and control plants the average sales in the municipality of the plant were around NOK 11 thousand, while the unemployment rate hovered around 2 percent.
Data on Statutory Payroll Tax Rates has been collected by Statistics Norway on the basis of administrative data sources. In figure 5 we observe the statutory tax rates. Most notably, in the year 2000 the tax rate for affected plants was equalized to the one in the advanced economic zone 1, until the year 2006. This effectively puts zone 1 as the control group in our identification strategy.

Figure 5: Statutory Payroll Tax Rates per Zone

4.3 Comparison of Means

We begin our analysis of the payroll tax reform with a comparison of mean wage rate and hours in Table 2. The first panel shows data for the wage rate split into four cells. The rows split the data into a pre-reform period called Baseline and reform period called Treatment, which starts from the year 2000 onwards. The columns split the data by treatment status of the plants. Each cell shows the mean of the log wage rate,
Table 2: DD Analysis of Mean Wage Rate and Hours

<table>
<thead>
<tr>
<th>Period</th>
<th>Treated Plants</th>
<th>Control plants</th>
<th>Difference</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel A. Wage Rate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (1996-1999)</td>
<td>4.93</td>
<td>5.00</td>
<td>-0.067***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.010)</td>
<td>(0.021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[516]</td>
<td>[1,562]</td>
<td>[2,078]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment(2000-2012)</td>
<td>5.27</td>
<td>5.39</td>
<td>-0.127***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.005)</td>
<td>(0.009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[1,950]</td>
<td>[4,608]</td>
<td>[6,558]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference over time</td>
<td>-0.33***</td>
<td>-0.39***</td>
<td>DD_{WR}=0.065***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[2,466]</td>
<td>[6,170]</td>
<td>[8,636]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B. Hours</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (1996-1999)</td>
<td>9.9</td>
<td>10.09</td>
<td>-0.19**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.085)</td>
<td>(0.047)</td>
<td>(0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[516]</td>
<td>[1,562]</td>
<td>[2,078]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment(2000-2012)</td>
<td>9.8</td>
<td>10.05</td>
<td>-0.25***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.027)</td>
<td>(0.046)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[1,950]</td>
<td>[4,608]</td>
<td>[6,558]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference over time</td>
<td>0.1</td>
<td>0.04</td>
<td>DD_{H}=0.12*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.085)</td>
<td>(0.054)</td>
<td>(0.061)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[2,466]</td>
<td>[6170]</td>
<td>[8,636]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Each cell shows the mean wage rate for panel A and the mean of hours for panel B. All variables are in logs. Robust standard errors in parenthesis. Asterisks denote: * * * p < 0.01, * * p < 0.05, * p < 0.1. Number of observations in square brackets.
the standard error in parenthesis and observations in square brackets.

The mean wage rate fell by 0.33 log points for treated plants in the treatment period versus a 0.39 decrease for control plants. Therefore, the wage rate declined by 0.065 log points more for control plants relative to treated plants. We see that control plants paid a higher wage rate throughout the sample period, with a relative increase in the gap between treated and control to 0.127 log points following the payroll tax reform. At the same time, in panel B we observe that the total hours worked at the plant did not decline significantly over time for both treated and control plants and the difference of roughly 0.2 log points between the two groups remained constant throughout the whole observation period.

### 4.4 Incidence

In this subsection we present results for the incidence and elasticities of the payroll tax in Norway. We start by establishing the incidence through estimating equation 18. We treat the incidence as the result of the first stage in an IV model. We proceed to recover labor demand elasticities by estimating equation 17 as the second stage. In table 3 we present the results on the incidence of the payroll tax. In panel A we estimate specification 18 by omitting the control variable in order to arrive at an indirect effect. In column 1 we find a baseline incidence parameter not significantly different than 0, which hints that the incidence falls on the employers, as the gross wage per hour does not seem to decline with a change in the payroll tax. As we add control variables in panel B, we evaluate the direct effect on the incidence, finding a coefficient of -0.43. Hence, when the payroll tax increases by 1 percent, gross wages decrease by 43 percent. This implies that 43 percent of the payroll tax is shifted on to employees, whereas 57 percent remains with the employer. This result is crucial to our application of the RER, as only by shared incidence can we pin down both shifts in the demand and the supply.

In column 2 and 3 we consider the effect of the payroll tax on hours worked and employees at the plant level. We observe that an increase in the tax leads to a decrease in both the hours and employees at the plant. The similarity in coefficients hints that the observed decrease in hours is driven by the decrease in employees at the plant and hours do not seem to be individually adjusted.

In panels C and D we consider the response to the payroll tax of two subsamples of firms. Firstly, we consider whether the incidence is different with respect to big plants where many workers are employed. We find that the incidence on employees is slightly higher than the average one in panel B, with 47 percent going to employees. Surprisingly, this leads also to a bigger response on the extensive margin, considering the number of employees employed in column 5. In the next subsection we discuss whether this effect is driven by demand or supply side market decisions and what it implies for the elasticity of labor demand.

Second, in panel D we consider a subsample of plants where the wage paid per worker is higher than the average for each tax zone - year observation. Here we observe that in this world of high rollers, the incidence shifted to employees is much higher - 60 percent, while the extensive margin effects are bigger than the average shown in panels A to C.

### 4.5 Labor Demand

In columns 1 and 2 of Table 4 we consider what do the adjustments on the extensive margin imply for the labor demand, by looking at the effect of an increase in the wage cost on hours and employees, where we use the payroll tax as an instrument in equation 17. By adding 1 to the coefficient in column 1 of table 3 we obtain the first stage of the IV regression underlying the labor demand specifications in columns 1-3.

In columns 4-6 we look at the labor supply responses, by making the RER assumption. The coefficients on Gross Wage Rate in columns 1 and 2 show the effect of an increase in the wage cost on the total hours worked at the plant and the number of employees. Column 3 considers an alternative specification when the dependent variable in the first stage is the Gross Wage Rate per employee. Similarly, columns 4 and 5 show the effect of an increase in the wage rate exclusive of the payroll tax and column 6 shows results for the effect of the wage rate per employee.
Table 3: The Effect of the Payroll Tax

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross Wage per Hour</td>
<td>Hours</td>
<td>Employees</td>
</tr>
<tr>
<td><strong>Panel A. Baseline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.283</td>
<td>-2.743***</td>
<td>-2.822***</td>
</tr>
<tr>
<td></td>
<td>(0.210)</td>
<td>(0.715)</td>
<td>(0.681)</td>
</tr>
<tr>
<td>Observations</td>
<td>8565</td>
<td>8565</td>
<td>8565</td>
</tr>
<tr>
<td><strong>Panel B. Direct Effect</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.438**</td>
<td>-3.470***</td>
<td>-3.488***</td>
</tr>
<tr>
<td></td>
<td>(0.208)</td>
<td>(0.695)</td>
<td>(0.663)</td>
</tr>
<tr>
<td>Observations</td>
<td>8563</td>
<td>8563</td>
<td>8563</td>
</tr>
<tr>
<td><strong>Panel C. Big Employers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.475**</td>
<td>-3.678***</td>
<td>-3.700***</td>
</tr>
<tr>
<td></td>
<td>(0.220)</td>
<td>(0.740)</td>
<td>(0.704)</td>
</tr>
<tr>
<td>Observations</td>
<td>4226</td>
<td>4226</td>
<td>4226</td>
</tr>
<tr>
<td><strong>Panel D. Big Wages</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.605**</td>
<td>-4.431***</td>
<td>-4.303***</td>
</tr>
<tr>
<td></td>
<td>(0.299)</td>
<td>(0.992)</td>
<td>(0.947)</td>
</tr>
<tr>
<td>Observations</td>
<td>4216</td>
<td>4216</td>
<td>4216</td>
</tr>
</tbody>
</table>

**Notes:** The dependent variable in each regression is shown at the top of the column. Statutory Payroll Tax is transformed as log(1 + Payroll Tax Rate). Regressions include plant × firm, sector and year fixed effects. The estimates are weighted by the number of employees at the plant. All regressions include as control variables the average sales and the unemployment rate at the municipality level. All variables are in logs. Robust standard errors. Asterisks denote: ***p < 0.01, **p < 0.05, *p < 0.1.
Looking at the baseline in panel A we observe elasticities of labor demand between -4 to -3.5, while the elasticities of labor supply are large positive numbers. This is in line with the incidence parameter in panel A of table 3. Given that the incidence falls on the employer, we do not have information to evaluate the effect of the wage rate on the labor supply.

Yet, as the incidence in panel B is shared between employers and employees we are able to identify a direct effect elasticity of -6 for labor demand and around 8 for labor supply. When we look at firms with many employees (panel C) we observe a smaller demand elasticity and similar supply elasticity. In firms with big wages the elasticity of labor demand attains values between -16 and -10, while the elasticity of labor supply decreases relatively in range to 5.8 to 7.1.

Robustness Checks. In the Appendix in Table A.1 we show several robustness checks to the main result on the incidence of the payroll tax. We saturate the estimated model with by adding sector time interacted fixed effect and absorbing more variation. The incidence appears to shift 46 percent of the burden to the employees, which is a similar estimate to the one in panel B in table 3. We replicate our results by applying a first difference estimator. We run a placebo test, by simulating the tax reforms for all other plants that were not affected by the tax reform. We find that in the absence of a tax reform the estimated coefficient on incidence is close to 0, while hours and employees tend to increase over the sample period, showing that our estimates on the extensive margin are rather an lower bound for the reaction of the manufacturing plants to such a payroll tax reform and the actual effect is higher. The elasticities for the labor demand and supply remain similar to the previous estimates except in the case of the placebo where we estimate zeros for labor demand.

In Table A.2 we look at the heterogeneity of the incidence and elasticities between the three sectors in our sample: steel production, mining and shipyards. We find that the incidence is fully shifted to employees in the steel production sector, allowing us to estimate elasticities of labor supply close to 0. Mining, which is the biggest sector of the 3, shows incidence and elasticities consistent with the main results. Shipyards show an incidence shifted to the employer, with a demand elasticity between -5.5 and -9.

In A.3 we present results for the service industry, exploiting the variation induced by the second tax reform in 2006. We find that the incidence of the payroll tax is shifted onto employers, followed by large decrease in the number of employees for the average firm. The fact that the incidence is shifted to employers results in a poor first stage for estimating the labor demand elasticity, which attains a value of -15.

Extension In Table 5 we extend our contribution by looking at the effect of the payroll tax on sales and gross investment within the plant. We look at the direct effect in panel A and we find that sales decrease by 4 percent and gross investment by 8 percent, albeit not significantly. When we look at big employers and big wage firms in panels B and C we find larger effects, and decreases of 5 to 8 percent on sales and 10 to 22 percent on gross investment. Unfortunately, in all cases the cases the estimates are plagues by a weak instrument bias. Given that around half of the incidence is shifted onto employees, there does not remain variation to estimate the effect of the wage rate on sales and gross investment.

5 Conclusion

References


Aaronson, Daniel, and Eric French (2009) ‘The effects of progressive taxation on labor supply when hours and wages are jointly determined.’ Journal of Human Resources 44(2), 386–408

### Table 4: The Effect of the Payroll Tax

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hours Employees</td>
<td>Hours Employees</td>
<td>Hours Employees</td>
<td>Hours Employees</td>
<td>Hours Employees</td>
<td>Hours Employees</td>
</tr>
<tr>
<td><strong>Panel A. Baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-3.824*** (1.345)</td>
<td>-3.935*** (1.437)</td>
<td>-3.541*** (1.226)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate per Employee Incl. Payroll Tax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.703 (8.108)</td>
<td>9.985 (7.985)</td>
</tr>
<tr>
<td>Gross Wage per Employee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13.909 (14.527)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>8565</td>
<td>8565</td>
<td>8565</td>
<td>8565</td>
<td>8565</td>
<td>8565</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>11.626</td>
<td>11.626</td>
<td>15.185</td>
<td>1.806</td>
<td>1.806</td>
<td>0.984</td>
</tr>
<tr>
<td><strong>Panel B. Direct Effect</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-6.175*** (2.362)</td>
<td>-6.207** (2.485)</td>
<td>-6.017*** (2.277)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate per Employee Incl. Payroll Tax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.923* (4.387)</td>
<td>7.964* (4.210)</td>
</tr>
<tr>
<td>Gross Wage per Employee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.299** (4.221)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>8563</td>
<td>8563</td>
<td>8563</td>
<td>8563</td>
<td>8563</td>
<td>8563</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>7.274</td>
<td>7.274</td>
<td>8.862</td>
<td>4.418</td>
<td>4.418</td>
<td>4.658</td>
</tr>
<tr>
<td><strong>Panel C. Big Employers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-7.006** (2.976)</td>
<td>-7.047*** (3.126)</td>
<td>-6.768** (2.805)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate per Employee Incl. Payroll Tax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.742* (4.208)</td>
<td>7.788* (4.036)</td>
</tr>
<tr>
<td>Gross Wage per Employee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.160** (4.081)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>4226</td>
<td>4226</td>
<td>4226</td>
<td>4226</td>
<td>4226</td>
<td>4226</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>5.678</td>
<td>5.678</td>
<td>7.051</td>
<td>4.650</td>
<td>4.650</td>
<td>4.849</td>
</tr>
<tr>
<td><strong>Panel D. Big Wages</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-11.224 (8.329)</td>
<td>-10.900 (8.405)</td>
<td>-16.125 (15.240)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate per Employee Incl. Payroll Tax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.322* (4.271)</td>
<td>7.110* (3.951)</td>
</tr>
<tr>
<td>Gross Wage per Employee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.870** (2.375)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>4216</td>
<td>4216</td>
<td>4216</td>
<td>4216</td>
<td>4216</td>
<td>4216</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>1.748</td>
<td>1.748</td>
<td>1.174</td>
<td>4.107</td>
<td>4.107</td>
<td>8.861</td>
</tr>
</tbody>
</table>

**Notes:** The dependent variable in each regression is shown at the top of the column. Statutory Payroll Tax is transformed as log(1 + Payroll Tax Rate). Regressions include plant \times firm, sector and year fixed effects. The estimates are weighted by the number of employees at the plant. Kleibergen-Paap F-statistic at the bottom. All regressions include as control variables the average sales and the unemployment rate at the municipality level. All variables are in logs. Robust standard errors. Asterisks denote: \* \* \* \ p < 0.01, \* \* \ p < 0.05, \* \ p < 0.1.
Table 5: The Effect of the Payroll Tax On Sales and Investment

<table>
<thead>
<tr>
<th></th>
<th>(1) Sales</th>
<th>(2) Gross Investment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Baseline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-4.339*</td>
<td>-8.361</td>
</tr>
<tr>
<td></td>
<td>(2.310)</td>
<td>(6.696)</td>
</tr>
<tr>
<td>Observations</td>
<td>8563</td>
<td>5664</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>7.274</td>
<td>3.703</td>
</tr>
<tr>
<td><strong>Panel B. Big Employers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-5.087*</td>
<td>-10.158</td>
</tr>
<tr>
<td></td>
<td>(2.848)</td>
<td>(8.430)</td>
</tr>
<tr>
<td>Observations</td>
<td>4226</td>
<td>3249</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>5.678</td>
<td>2.845</td>
</tr>
<tr>
<td><strong>Panel C. Big Wages</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-8.763</td>
<td>-22.527</td>
</tr>
<tr>
<td></td>
<td>(7.418)</td>
<td>(42.709)</td>
</tr>
<tr>
<td>Observations</td>
<td>4216</td>
<td>2823</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>1.748</td>
<td>0.323</td>
</tr>
</tbody>
</table>

Notes: The dependent variable in each regression is shown at the top of the column. Statutory Payroll Tax is transformed as $\log(1 + \text{Payroll Tax Rate})$. Regressions include plant $\times$ firm, sector and year fixed effects. The estimates are weighted by the number of employees at the plant. Kleibergen-Paap F-statistic at the bottom. All regressions include as control variables the average sales and the unemployment rate at the municipality level. All variables are in logs. Robust standard errors. Asterisks denote: ** ** $p < 0.01$, * * $p < 0.05$, * $p < 0.1$. 
mandates with an application to the us unemployment insurance program.’ *Journal of Public Economics* 65(2), 119–145


Doerrenberg, Philipp, Andreas Peichl, and Sebastian Siegloch (forthcoming) ‘The elasticity of taxable income in the presence of deduction possibilities.’ *Journal of Public Economics*


25


Johansen, Frode, and Tor Jakob Klette (1997) ‘Wage and employment effects of payroll taxes and investment subsidies.’ *mimeo*


Appendix A

Table A.1: Robustness Checks

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross Wage per Hour</td>
<td>Hours</td>
<td>Employees</td>
<td>IV Hours</td>
<td>Employees</td>
<td>IV Hours</td>
<td>Employees</td>
<td>IV Hours</td>
<td>Employees</td>
</tr>
<tr>
<td>Panel A: With SectorxTime f.e.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.462**</td>
<td>-2.357***</td>
<td>-2.718***</td>
<td>(0.193)</td>
<td>(0.650)</td>
<td>(0.631)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-4.383**</td>
<td>-5.055**</td>
<td>(1.805)</td>
<td>(2.083)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate per Employee Incl. Payroll Tax</td>
<td>-3.024***</td>
<td>(0.900)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage per Employee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>8522</td>
<td>8522</td>
<td>8522</td>
<td>8522</td>
<td>8522</td>
<td>8522</td>
<td>8522</td>
<td>8522</td>
<td>8522</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>7.741</td>
<td>7.741</td>
<td>23.925</td>
<td>5.725</td>
<td>5.725</td>
<td>0.304</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B: First Differences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S.Statutory Payroll Tax</td>
<td>-0.418</td>
<td>2.740***</td>
<td>2.037**</td>
<td>(0.434)</td>
<td>(0.980)</td>
<td>(0.964)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S.Gross Wage Rate per Employee Incl. Payroll Tax</td>
<td>1.585*</td>
<td>(0.934)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S.Wage Rate</td>
<td></td>
<td>-7.370</td>
<td>-5.505</td>
<td>(7.281)</td>
<td>(6.122)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S.Gross Wage per Employee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>7058</td>
<td>7058</td>
<td>7058</td>
<td>7058</td>
<td>7058</td>
<td>6819</td>
<td>6819</td>
<td>6819</td>
<td></td>
</tr>
<tr>
<td>F-Statistic</td>
<td>1.792</td>
<td>1.792</td>
<td>16.546</td>
<td>0.879</td>
<td>0.879</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel C: Placebo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.040</td>
<td>-0.312*</td>
<td></td>
<td>(0.057)</td>
<td>(0.169)</td>
<td>(0.168)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-0.353**</td>
<td>-0.325*</td>
<td>(0.173)</td>
<td>(0.174)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage Rate per Employee Incl. Payroll Tax</td>
<td>-0.335*</td>
<td>(0.177)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Wage per Employee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>165553</td>
<td>165553</td>
<td>165553</td>
<td>165553</td>
<td>165553</td>
<td>165553</td>
<td>165553</td>
<td>165553</td>
<td>165553</td>
</tr>
</tbody>
</table>

Notes: The dependent variable in each regression is shown at the top of the column. Statutory Payroll Tax is transformed as log(1 + Payroll Tax Rate). Regressions include plant × firm, sector and year fixed effects except in panel A and B. The estimates are weighted by the number of employees at the plant. The last two columns present estimates from IV regressions, with Kleibergen-Paap F-statistic at the bottom. Wage Cost as independent variable is the gross wage per hour plus the statutory payroll tax. All regressions include as control variables the average sales and the unemployment rate at the municipality level. All variables are in logs. Robust standard errors. Asterisks denote: ∗ ∗ ∗ p < 0.01, ∗ ∗ p < 0.05, ∗ p < 0.1.
Table A.2: The Results by Sector

<table>
<thead>
<tr>
<th>Panel A. Steel Production</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.947**</td>
<td>0.170</td>
<td>0.768</td>
<td>(0.482)</td>
<td>(1.759)</td>
<td>(1.805)</td>
<td>0.261</td>
<td>(2.729)</td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>3.191</td>
<td>14.417</td>
<td>(46.611)</td>
<td>14.17</td>
<td>(143.062)</td>
<td>0.012</td>
<td>(1.842)</td>
<td>(1.849)</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>0.012</td>
<td>2.898</td>
<td>3.857</td>
<td>0.831</td>
<td>0.012</td>
<td>2.898</td>
<td>3.857</td>
<td>0.831</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Mining</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.442**</td>
<td>-1.874***</td>
<td>-1.548***</td>
<td>(0.220)</td>
<td>(0.589)</td>
<td>(0.592)</td>
<td>-0.488</td>
<td>(4.992)</td>
</tr>
<tr>
<td>Gross Wage Rate Incl. Payroll Tax</td>
<td>-3.357**</td>
<td>-2.773*</td>
<td>(1.683)</td>
<td>(1.420)</td>
<td>-2.119***</td>
<td>(0.801)</td>
<td>3.503</td>
<td>(2.331)</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>6.451</td>
<td>15.917</td>
<td>4.040</td>
<td>0.272</td>
<td>6.451</td>
<td>15.917</td>
<td>4.040</td>
<td>0.272</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C. Shipyards</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.420</td>
<td>-5.430***</td>
<td>-5.041***</td>
<td>(0.379)</td>
<td>(1.116)</td>
<td>(1.118)</td>
<td>16.208</td>
<td>(32.082)</td>
</tr>
<tr>
<td>F-Statistic</td>
<td>12.015</td>
<td>12.944</td>
<td>182.061</td>
<td>(2758.562)</td>
<td>4.617</td>
<td>1.224</td>
<td>1.224</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Notes: The dependent variable in each regression is shown at the top of the column. Statutory Payroll Tax is transformed as log(1 + Payroll Tax Rate). Regressions include plant × firm, sector and year fixed effects except in panel A and B. The estimates are weighted by the number of employees at the plant. The last two columns present estimates from IV regressions, with Kleibergen-Paap F-statistic at the bottom. Wage Cost as independent variable is the gross wage per hour plus the statutory payroll tax. All regressions include as control variables the average sales and the unemployment rate at the municipality level. All variables are in logs. Robust standard errors. Asterisks denote: ** p < 0.01, * p < 0.05, * p < 0.1.

Table A.3: Services, year > 2002

<table>
<thead>
<tr>
<th>Panel</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statutory Payroll Tax</td>
<td>-0.143</td>
<td>-1.252</td>
<td>-1.341**</td>
<td>(0.517)</td>
</tr>
<tr>
<td>Wage per Employee</td>
<td>-15.125</td>
<td>(113.630)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-Statistic</td>
<td>2.342</td>
<td>2.342</td>
<td>4.617</td>
<td>1.224</td>
</tr>
</tbody>
</table>

Notes: The dependent variable in each regression is shown at the top of the column. Statutory Payroll Tax is transformed as log(1 + Payroll Tax Rate). Regressions include firm, sector and year fixed effects. The estimates in the first column are weighted by the number of employees at the plant. The last column present estimates from IV regressions, with Kleibergen-Paap F-statistic at the bottom. Wage Cost as independent variable is the gross wage per hour plus the statutory payroll tax. All regressions include as control variables the average sales and the unemployment rate at the municipality level. All variables are in logs. Robust standard errors. Asterisks denote: ** p < 0.01, * p < 0.05, * p < 0.1. The underlying sample runs from 2002 to 2012.