Paper presented at the IZA/World Bank Workshop: Institutions and Informal Employment in Emerging and Transition Economies, IZA, Bonn, 9-11 June 2011.

An attempt to identify grey employment – Estimation of wage under-reporting and tests of the predictions

Péter Elek[a] - János Köllő[b] – Balázs Reizer[c] - Péter A. Szabó[d]

27 May 2011[1]

Abstract. We estimate a double hurdle (DH) model of the Hungarian wage distribution assuming truncation at the minimum wage and wage under-reporting (i.e. compensation consisting of the minimum wage, subject to taxation, and an unreported cash supplement). We classify individuals and firms as 'cheaters' and 'non-cheaters' on the basis of the DH estimates and check how they reacted to the doubling of the minimum wage in 2001-2002 and to the introduction of a minimum contribution base, equal to 200 per cent of the minimum wage, in 2007. We find evidence that cheaters and non-cheaters responded differently to the treatments under investigation. First, minimum wage effects significantly differed across the two groups of firms affected by the large hikes. Second, the MW workers employed by suspected cheaters (as of 2006) experienced faster wage growth and/or were more likely to earn 200 per cent of the minimum wage in 2007 than their counterparts in non-cheating firms. The results suggest that the DH model is able to locate wage under-reporting with acceptable precision.

1. Introduction

The evasion of payroll taxes has two main forms. One is unreported (black) employment, when the employee is not registered and neither she nor her employer pays any taxes. The other main form is the under-reporting of wages (grey employment); when the compensation consists of an officially paid amount, subject to taxation, and an unreported cash supplement also known as an "envelope wage" or "under the counter payment". In order to maximize the total evaded tax, the officially paid wage is often (but not always) chosen as the minimum wage (MW).

In this paper we estimate the prevalence of disguised MW earners with the double hurdle (DH) econometric model (first proposed by Cragg 1971) using linked employer-

a) Eötvös Loránd University, Budapest b) Institute of Economics, Budapest c) Central European University d) Reformed Presbyterian Church of Central and Eastern Europe
[1] Corresponding author: János Köllő (IE and IZA), kollo@econ.core.hu

employee data. Our application of the DH model assumes that a spike at the MW is observed for two reasons (i) because of constraints and costs preventing firms from firing low-productivity workers and (ii) because of tax fraud. Consequently, a worker's genuine wage is observed only if her productivity exceeds the MW *and* her wage is fully reported. The DH model simultaneously deals with the truncation problem and selection to tax fraud, and estimates the probability of cheating for each MW earner. In the possession of the parameters one can also simulate the 'genuine' wages of MW earners.

The DH model's reliance on distributional properties (as well as the difficulties of finding the selection equation's own regressors) warns us not to take the estimates at face value even though they seem sensible *per se*. Therefore, we test the validity of the DH results in two ways.

Test 1 looks at the aftermaths of a particularly large hike in the minimum wage in 2001-2002, using a panel of small firms. We classify firms as cheaters and non-cheaters on the basis of the DH results and look at how their exposure to the MW hikes affected their wages and employment in 2000-2003. The expectation is that in the group of firms suspected of tax evasion the base-period share of MW earners had smaller positive impact on the observed rate of wage growth, weaker negative effect on actual labor costs (and therefore employment) and cheaters also had weaker incentive to substitute skilled for unskilled labor. We test these assumptions by allowing exposure to the MW hike to have different impact on cheaters and non-cheaters in a variety of model specifications. In most of the specifications the results significantly differ across the groups formed on the basis of the DH results.

Test 2 examines the introduction of a minimum contribution base amounting to 200 per cent of the minimum wage (2MW), in 2007. After the introduction of the reform, firms paying wages lower than 2MW faced a high probability of tax authority audit and a high risk of being detected as cheaters. The reform created incentives for cheating firms to raise the reported wages of MW earners while non-cheaters (those paying genuine minimum wages) had no interest to do so. We distinguish cheaters from non-cheaters on the basis of the DH results and check how the cheating proxies affected the probability that a worker earning the MW in 2006 earned 2MW in 2007. We also study how the

wages of MW earners of suspected cheating and non-cheating firms changed in 2006-2007. In most of the specifications we find that the MW earners of suspected cheaters were more likely to be paid 2MW in 2007 and/or experienced faster wage growth than their observationally similar counterparts in non-cheating firms. Furthermore, we find that employment in cheating firms was adversely affected by the reform.

At least in the East and South-East of Europe, MW policies are strongly influenced by the conviction that nearly all MW workers earn untaxed side payments. Our results suggest that while the suspicions are not groundless they are strongly overstated. We estimate the share of 'disguised' MW earners to be between 35 and 55 per cent among all MW earners. Non-cheating businesses account for a sizeable minority (30-40 per cent) even among small, low-wage firms employing MW workers. Furthermore, our results yield supplementary evidence on the adverse employment effects of Hungary's unconventional MW policies (see Halpern et al. 2004, Kertesi and Köllő 2003), at least in the case of non-cheating firms. The statistical profiles derivable from the DH model may help the better targeting of tax authority inspection and facilitate more circumspect MW policies.

The paper is organized as follows. Section 2 gives a brief overview of the literature while Section 3 introduces the two 'natural experiments' that we study in detail. Section 4 introduces the DH model, explains the estimation of its parameters and shows how the probability of cheating and the 'genuine' wage can be simulated for each MW earner. It discusses how we classify workers and firms on the basis of the DH estimates and explains how the two tests proceed. Section 5 introduces the data. Section 6 presents the estimates of the DH model and the results of the two tests. Finally, Section 7 concludes.

2. Wage under-reporting and the minimum wage – An under-researched area

Compared to the vast literature on income under-reporting and MW regulations, respectively, the body of research on how these two areas relate to each other seems rather thin. Most of what we know empirically about this relationship comes from anecdotal evidence, inspection of aggregate data, scarce survey results and a few

attempts to identify the incidence of envelope wages indirectly. Theoretical work is largely missing.

Cross-country data suggest a positive correlation between the size of the spike of the wage distribution at the MW and the estimated size of the informal economy (Tonin 2006). Several accession countries including Hungary, Latvia, Lithuania and Romania have (or had) high shares of MW earners, while their Kaitz-indices are (were) in the middle range, suggesting that disguised MWs may be particularly wide-spread in these countries. Similar observations are interpreted in a similar way in World Bank (2005).

Erdogdu (2008) reports on the basis of several surveys that under the counter payments are prevalent in the wage policy of Turkish firms. There is a relatively extensive literature focusing on grey employment in the Baltic states. Masso and Krillo (2008) point out that 16-23 percent of the MW earners received envelope wages in Estonia and Latvia but only 8 percent in Lithuania in 1998. Meriküll and Staehr (2008) show that young employees and people working in construction and trade are most likely to get unreported cash supplement on top of their official salary in the three Baltic states. Kris et al. (2007) present similar results on the distribution of envelope wages using three different Estonian data sets. According to the Eurobarometer survey conducted by the European Commission in 2007 (European Commission 2007), 5 per cent of employees in the EU receive part or all of their regular income untaxed and this ratio is over 10 per cent in some central and eastern European countries (8 per cent in Hungary) but there is no information on how many of them are officially paid the MW.

Some studies obtain evidence on disguised MWs indirectly, by comparing the reported consumption-income profiles of households. Using household budget survey data from Hungary, Benedek et al (2006) looked at the winners and losers from the 2001-2002 MW hikes. They observed income loss *without* the loss of a wage earner in the high-income brackets where substantial under-reporting is most likely to occur. For these households the increasing MW may have implied higher taxes and lower net income. Based on the same data set, Tonin (2007) analyzed changes in the food consumption of households affected by the 2000-2001 minimum wage hike compared to unaffected households of similar income. He found that food consumption fell in the treatment

group relative to the controls – a fact potentially explained by a fall in their net income as the reported part of their wages (and the associated tax burden) rose.

The theories of wage under-reporting (Allingham-Sandmo 1972, Yaniv 1988) shed light on the incentives to engage in tax fraud under alternative penalty and withdrawal schemes but they do not explicitly discuss the case of reporting the MW to tax authorities. This is the cost-minimizing choice for the firm (unless MW payment provokes audits thereby decreasing the expected gain from cheating) but it also requires the cooperation of workers. As Madzharova (2010) notes, if the actual or perceived linkages between contribution payments and pensions or access to health services are weak and/or workers see that their payments feed corruption rather than used to finance public services, they will be willing to accept the lowest possible reported wage.

The only theoretical model to our knowledge, which explicitly addresses the issue of wage under-reporting *cum* MW regulations, is Tonin (2006). Tonin argues that the MW induces some workers whose productivity is *above* the MW, but who would have declared less if there was no MW, to increase their declared earnings to the MW level. Workers with productivity below the MW either work in the black market or withdraw from the labor force while high-productivity workers are unaffected. This is a possible explanation of why a spike at the MW appears in the distribution of declared earnings.

While our empirical work is also concerned with the spike at the MW we do not maintain the assumption that the marginal products of those at the spike necessarily exceed the MW. Tonin's model looks at equilibrium in a perfectly competitive market while we look at episodes, when the MW (or its tax burden) was drastically increased. When the plan of increasing the MW to Ft 50,000 was announced, in 2000, 32.7 per cent of the private sector employees earned less than that. When the idea of the minimum contribution base came up, in 2006, 58 per cent of the private sector workers had wages below 2MW. It is quite obvious that the vast majority of these workers remained in employment for a protracted period (or until recently) while many of them had productivity below the aforementioned thresholds just after the hikes. It took time until mobility between jobs, changes of the product mix and technology, adult training and other forms of adjustment could restore (if at all) the optimum condition for mutually

gainful employment without causing massive unemployment in between. Therefore we stick to the assumption that in the periods under examination the spike at the MW was explained by under-reporting *and* the continuing employment of low-productivity workers – two different phenomena that we try to model following the DH approach.

3. The minimum wage and the wage distribution in Hungary

MW regulations played increasingly important role in shaping the Hungarian wage distribution. As shown in Figure 1, the distribution was close to lognormal in 1986, three years before the fall of state socialism and the introduction of MW regulations. In the next ten years a minor spike at the MW developed, which became more enhanced between 1995 and 2000, despite the fact that the Kaitz-index fell from 45 to 36 per cent in this period.[2]

Figure 1

In 2001–2002 the MW was nearly doubled in nominal terms, resulting in a 14 percentage point rise in the Kaitz-index and a 70.5 per cent increase in the real value of the gross minimum wage.[3] The wage distribution was severely deformed by the hikes, with the fraction of private sector employees earning the MW ±5 per cent jumping from 5.3 per cent in May 2000 to 18.1 per cent in May 2002.[4]

The distribution preserved its shape until 2007, when a second spike appeared at 200 per cent of the MW.[5] That year, the Hungarian government introduced a minimum social security contribution base amounting to 2MW. Firms were allowed to pay wages lower than 2MW but in case they did so they faced a high probability of tax authority audit and

[2] Unless otherwise indicated, all wage data are based on the Wage Survey introduced in Section 5.

[3] The minimum wage increased from Ft 25,500 in 2000 to Ft 40,000 on January 1, 2001 and Ft 50,000 on January 1, 2002. See Kertesi and Köllő (2003) on the motives and aftermaths of the large hikes.

[4] The data relate to firms employing more than 5 workers and come from the Wage Survey (see later).

[5] In the same time further minima were introduced for young and older skilled workers (1.2MW, 1.25MW) that flattened the spike near the MW.

a high risk of being detected as cheaters (for paying disguised MW or for other reasons).[6]

The suspicion that the crowding of workers at the MW and 2MW is partly explained by wage under-reporting is difficult to avert. In 2003, shortly after the MW was doubled, the highest fraction of MW earners (27 per cent) was found among small firm managers, and about one in ten top managers in larger firms also earned the MW. High shares could be observed in a number of freelance occupations such as architects, lawyers, accountants, business and tax advisors, agents, brokers, artists, writers, film-makers, actors and musicians (15-17 per cent). The fraction was high in personal services (22 per cent), trade (25 per cent) and house-building - sectors where cash transactions with customers frequently occur. In several low-wage occupations the fraction earning the MW fell short of the above-mentioned levels (Table 1).

Table 1

Further doubts arise if we look at the wage distribution within occupations. The panels of Figure 2 show the shape of the density functions for three occupations. In the case of engineers and natural scientists the MW plays a minor role: the wage distribution is close to the lognormal. In the case of unskilled laborers and casual workers the distribution is strongly skewed at the MW with a small number of workers earning substantially more than that. By contrast, the wage distribution of top managers has a bimodal shape with one mode located at the MW and another at 500 per cent of the MW. This shape rather clearly points to a minority of managers under-reporting their wages.[7]

Figure 2

With the help of the double hurdle model we can utilize the information content of the different shapes of the wage distributions. In the next section we summarize how the estimation proceeds, how the probability of under-reporting and the MW earners' 'genuine' wages are derived, and how we test the validity of the estimates.

---

[6] Similar minimum contribution levels were introduced in Bulgaria and Croatia in 2003. The Hungarian regulations remained in effect until 2011.

[7] See Appendix 2 for further details of Hungary's MW regulations.

4. Estimation methods

4.1. The double hurdle model

Let us use the notation *y* for the logarithm of the "true" wage, i.e. of the wage which would prevail in the absence of MW and under-reporting. *y* is determined by some characteristics *X* of the employee and the firm, and we assume that its distribution is conditionally normal with expectation $X\beta$ and variance σ^2. (This is a standard assumption in the literature; see e.g. Meyer and Wise 1983a and 1983b.) In the presence of MW and under-reporting, a spike appears at the MW in the wage distribution. The observed wage (the logarithm of which will be denoted by *y**) may be equal to the MW for two reasons: because of constraints and costs preventing firms from firing low-productivity workers (in the simplest case those whose genuine wage would fall below the MW), or because of tax fraud (when the MW is reported to the authorities but an unobserved cash supplement is also given). The probability of cheating is determined by some characteristics *Z* of the employee and the firm, and *X* may be different from *Z*. Formally, omitting subscript *i* for the individual, the following model governs *y* and *y**:

(1) $$y = X\beta + u\,,$$

and we observe the reported wage *y** according to the rule:

(2) $$y^* = \begin{cases} y & \text{if } X\beta + u > m \text{ and } Z\gamma + v > 0 \\ m & \text{otherwise} \end{cases},$$

where *m* is the logarithm of the MW. Under-reporting occurs when both $X\beta+u>m$ and $Z\gamma+v\leq 0$ hold, and in this case the observed wage is equal to the MW. The residuals *u* and *v* are zero-mean normally distributed, possibly correlated (ρ) random variables. σ^2 stands for the variance of *u* while the variance of *v* is set equal to unity without loss of generality, hence the covariance matrix of (*u*, *v*) is given by:

(3) $$S = \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & 1 \end{pmatrix}.$$

This is the double hurdle model first proposed by Cragg (1971), with the restriction ρ=0, to model the purchase of consumer goods in a setting where the decision to buy and the decision of how much to buy are governed by different processes. The name of the model comes from the fact that the spike of the distribution (in our case at the MW) is determined by two "hurdles": a standard tobit-type constraint (in our case following from the wage equation: $X\beta+u\leq m$) and a different second hurdle (following from the selection equation: $Z\gamma+v\leq 0$). Note that the standard tobit model is obtained as a special case when the second hurdle is not effective, e.g. when Z contains a sufficiently large constant and all other terms in γ are zero, or when X=Z, β=γ (apart from a constant), ρ=1 and σ=1. In our case, a second hurdle is needed because under-reporting and wage determination are governed by partly different processes.

Since the paper of Cragg the model and its extensions have been widely used to analyze consumer and producer behavior as well as problems in environmental and agricultural economics and banking (e.g. Labeaga 1999, Martinez-Espineira 2006, Moffatt 2005, Saz-Salazar and Rausell-Köster 2006, Teklewold et al. 2006). However, to our knowledge, only Shelkova (2007) used the model to analyze wage distributions, albeit in a setting different from ours.

Figure 3

In our application, the baseline DH model (1)-(3) has to be slightly modified in order to better capture the features of the wage formation process. The first problem to be addressed is that the log wage distribution is *not* truncated normal because of the crowding of wage earners just above the MW [8] (see Panel A in Figure 3). While at and above the median the distribution is close to the normal we have more workers on the left tail than expected under normality. This poses a problem because – as usual for nonlinear models – maximum likelihood estimation of the DH model yields consistent results only if the underlying distributions are well-specified. Therefore we apply a preliminary transformation that is roughly the identity at higher wages and accounts for

[8] This is explained by spillover effect as argued in Dickens et al. (1994) and elsewehere.

'crowding' at lower wages. We assume that instead of y^* we observe $g(y^*)$, where r is a coefficient to be estimated:

$$(4) \qquad g(x) = x + r \cdot \exp(-(x - m + r)/r) \text{ if } x \geq m - r.$$

Figure 4 shows the shape of g for the parameter choice r=0.3. By the preliminary transformation g^{-1} we can ensure that y^* is (truncated) normally distributed and hence the DH model can be applied. Our approach is in line with the double hurdle literature, where a preliminary transformation is often needed to achieve normality: Martinez-Espineira (2006) and Moffatt (2005) use the Box-Cox, while Yen and Jones (1997) apply the inverse hyperbolic sine transformation.

Figure 4

The second possible problem concerns our assumption that cheating employers report the MW (and not a larger wage) to the authorities. This is a reasonable assumption for 2003-2006 because firms could maximize the evaded tax this way and the chance of tax audit was not increased for MW-reporting firms before 2007. The model can be extended to allow for cheating above the MW (see Elek et al. 2009) but external (e.g. survey-based) information is needed to identify its parameters. In this paper, we stick to the simpler formulation.

4.2. Parameter estimation

First, the parameter r of the preliminary transformation (4) is estimated using two methods. The first method is purely statistical: wages above the MW are modeled with a distribution of shape $g(U)$ where U is truncated normal with expectation e and variance s. (The parameter vector (r,e,s) is estimated by maximum likelihood.) In the second method we create a quasi panel subsample of the LEED data for 2000-2002 and assign the median of the 2002 wages to the median of the 2000 wages of each percentile of the wage distribution in 2000. (See section 5 for details of the data set.) Then the wage-wage graph obtained this way is adjusted for the average wage growth in 2000-2002, and the function g (with unknown parameter r) is fitted to it with nonlinear least squares. Both methods yield r=0.32, thus the transformation g^{-1} is applied with this parameter value to

the observed log wages. Panel B in Figure 4 shows that the transformed log wages above the MW are approximately truncated normal. In what follows, we refer to the transformed log wage as *y**.

Using the properties of the conditional distributions of the bivariate normal distribution, the likelihood function of the DH model (1)-(3) can be shown to have the following form (for the sake of clarity, here we use subscripts *i* for the individuals):

$$(5) \qquad L = \prod_{y_i^*=m} \left[1-\Phi_{\rho,\sigma,1}\left(x_i\beta-m, z_i\gamma\right)\right] \cdot \prod_{y_i^*>m} \left[\Phi\left(\frac{z_i\gamma+\frac{\rho}{\sigma}\left(y_i^*-x_i\beta-m\right)}{\sqrt{1-\rho^2}} \right) \frac{1}{\sigma}\phi\left(\frac{y_i^*-x_i\beta-m}{\sigma} \right) \right],$$

where $\Phi_{\rho,\sigma,1}$ denotes the bivariate normal distribution with covariance matrix given in (3), while Φ and ϕ stand for the univariate standard normal distribution and density, respectively. Parameter estimation can be carried out with maximum likelihood, which is consistent and asymptotically normal if the underlying distributions are correctly specified.

4.3. Simulation of under-reporting probabilities and 'genuine' wages

In the possession of the parameters the probability of cheating for each MW earner can be estimated as:

$$(6) \qquad \begin{aligned} P(\text{underreporting}) &= P(X\beta+u>m, Z\gamma+v\le 0 \mid y^*=m) = \\ &= \frac{P(u>-X\beta+m)-P(u>-X\beta+m, v>-Z\gamma)}{1-P(u>-X\beta+m, v>-Z\gamma)} = \\ &= \frac{\Phi((X\beta-m)/\sigma)-\Phi_{\rho,\sigma,1}(X\beta-m, Z\gamma)}{1-\Phi_{\rho,\sigma,1}(X\beta-m, Z\gamma)} \end{aligned}$$

As a final step, we simulate the genuine wage of each MW earner as follows. We generate independent copies of bivariate normal random variables (*u*, *v*) with covariance matrix given in (3), and accept max (*Xβ+u*, *m*) as the genuine log-wage of an MW earner if *Xβ+u≤m* or *Zγ+v≤0*. If none of these conditions hold, the person cannot earn MW

according to the model. Technically, for each MW-earner, the (u,v) variables are simulated until at least one condition holds.

4.4. Classification of workers and firms on the basis of the DH estimates

Denoting the estimated probability of under-reporting by a MW earner with P and the simulated wage with w (i.e. $w=e^y$) in the following sections cheating behavior is assumed in case of (i) w>MW (ii) P>0.5 and (iii) w>1.5MW, respectively.[9] (Strictly speaking, on the individual level w>MW is equivalent to cheating but we use the other definitions as well to check robustness.) If we find at least one MW earner classified as "cheater" in a firm we treat the firm as a cheater.

4.5. Estimating the differential impacts of the large MW hike

The MW hike of 2001-2002 increased the share of MW earners by 14 percentage points in the private sector. Presumably, a part of the firms paid envelope wages on top of the MW already in 2000 but this practice clearly became prevalent after the large hikes. In either case we expect that exposure to the MW hike exerted stronger effect on the labor costs of non-cheaters.

A MW hike forces firms to raise the wages of affected workers (those earning less than the new minimum wage before the hike, or $MW_0 \leq w_0 < MW_1$) to or above MW_1. In at least the medium run, identical workers have to be paid the same net wage and it is reasonable to assume that the surplus of the new wage over MW_1 will be proportional to the surplus of the old wage over MW_0. Formally, we can write $w_1 = MW_1 + \rho(w_0 - MW_0)$, where $0 \leq \rho \leq 1$ is a spillover parameter. 'Never cheaters' fully pay their contributions (τw) while 'always cheaters' and new cheaters only pay after MW_0 and/or MW_1. That said, it is straightforward to see that the difference between never cheaters and always cheaters in terms of cost increase for identical workers will be $D=\rho\tau(w_0 - MW_0)>0$ and the difference between never cheaters and new cheaters will be $D=\rho\tau(w_0 - MW_0)\geq 0$. (Note that for always cheaters $w_0 > MW_0$, for new cheaters $w_0 \geq MW_0$ while for both new and

[9] We estimated our models applying further criteria (w>1.1 MW and w>2MW) but it did not alter the qualitative conclusions.

always cheaters $w_1>MW_1$ by definition). The higher the worker's starting wage, the stronger the spillover effect and the higher the social security contribution rate, the larger the difference between cheaters and non-cheaters in terms of cost increase implied by the MW hike.[10]

The difference between cheaters and non-cheaters (in terms of potential cost increase implied by the MW hike) is not constant: it is zero for firms with no affected employees and can be substantial if all of their employees are affected (i.e. earn less than MW_1). Therefore, we study the responses of suspected cheaters and non-cheaters group by group rather than simply including a 'cheater' dummy to the model.

Practically, we regress firm-level outcome variables (changes of reported wages, employment and the share of unskilled employees) on proxies of exposure to the large MW hike, and allow the exposure proxies to have different effects in the cheater and non-cheater groups. We do not split the sample into two parts; instead we interact the cheater/non cheater dummies with the proxies of exposure as in equation (7), where the log change of the observed average wage is on the left hand side:

$$(7)\ \Delta \ln w^* = \beta_1(Exposure \times Cheater) + \beta_2(Exposure \times Non\_cheater) + Z\gamma + u$$

In the equation w^* stands for the firm's average official (reported) wage (i.e. the average of the $w^*=e^{y^*}$ reported wages of the employees), Z comprises base-period firm characteristics and Δ stands for change between May 2000 and May 2003. Since the composition of employment might have changed in the observed period we also estimate the equation with residual wages on the left hand. Residual wages are calculated as the firm-level means of residuals from standard Mincer equations estimated using individual data from 2000 and 2003 (see section 5 for details). We also estimate equations similar to (7) with the log change of employment and the percentage

[10] In the case of some workers employed by 'always cheaters' the starting wage might have exceeded MW_1 while the reported wage was equal to MW_0 and increased to MW_1 after the hike. On the basis of previously quoted results from Benedek et al. (2006) and Tonin (2007), we hypothetize that actual net wages fell in these cases, partly offsetting the cost increase amounting to $\tau\Delta MW$. We believe that given the size of the MW hike, $w_0>MW_1$ infrequently occurred in 2000.

points change in the share of unskilled workers on the left hand. In each equation we expect $|\beta_1| < |\beta_2|$.

Exposure to the MW hike is measured with the base-period share of low-wage workers (those earning less than MW_1) and an indicator approximating average wage increase required to accommodate the MW hike under full compliance and no spillover effects. This 'MW shock' indicator, similar to that proposed by Machin et al. (2003), can be written as:

$$(8) \quad \omega = \ln \frac{MW_1 F + \overline{w}_H (1 - F)}{\overline{w}_F F + \overline{w}_H (1 - F)}$$

where MW_1 stands for the new minimum wage, F denotes the fraction of workers affected by the hike (those earning less than MW_1) and $\overline{w}_F$ and $\overline{w}_H$ denote the average observed wages of affected and unaffected (low-wage and high-wage) workers. We estimate specifications, where F and ω are defined using MW_1=40,000 and MW_1=50,000 (the MWs after the first and the second hike) as thresholds, respectively,

Suspected cheaters are identified on the basis of the DH model estimated for 2003 (not 2000). We opted for this solution because the available data and anecdotal evidence suggested that wage under-reporting became wide-spread during and after the large hikes. Since tax evasion helped to contain the growth of labor costs of both 'always cheaters' and 'new cheaters', we make no attempt to separate them.

4.6. Estimating the effects of the minimum contribution base

In 2007-2010, when the minimum contribution base was in effect, firms were expected to pay social security contribution amounting to at least 2τMW per worker (denoting the contribution rate with τ). Firms had the possibility to pay less than 2τMW provided that they presented convincing evidence that the going market wage for the workers in

question fell short of 2MW. These firms, however, faced a higher than average risk of tax authority audit. [11]

This system created incentives to raise the reported wages of 'disguised MW earners'. First, cheating firms could fully avert the risk of audit by paying 2MW instead of MW to their high-wage employees. Second, the reform gave a clear signal that the tax authority treated MW payment as a indicator of tax evasion. Therefore, by shifting MW earners to somewhere between MW and 2MW fraudulent firms could reduce the risk of audit. In the same time, non-cheating firms paying 'genuine' MW had less incentive to raise the wages of their MW earners.

We test if cheating and non-cheating firms reacted differently to the reform by estimating probit (9) and linear regression (10) for a quasi-panel of individuals observed in May 2006 and May 2007 (see section 5 for the details):

$$(9)\ P(w_1^* = 2MW_1 \mid w_0^* = MW_0) = \Phi(\beta[Classified\ \ as\ cheater] + Z\gamma)$$

$$(10)\ \Delta w^* \mid (w_0^* = MW_0) = \beta[Classified\ \ as\ cheater] + Z\gamma + u$$

where Z comprises worker and firm characteristics, MW_0 and MW_1 stand for the minimum wage in 2006 and 2007 and Δ denotes change in Forints. (Since equation (10) is estimated for the MW earners of 2006 the results also capture the effects in terms of percentage change.) In both equations the expectation is $\beta>0$.

Raising the reported wages of disguised MW earners implied costs for cheating firms; therefore we expect that their output and level of employment fell. We test these expectations by estimating firm-level regressions with log changes of employment and sales revenues on the left hand and a 'cheater' dummy included in the right hand side of the equations.

[11] We can regard this system as a version of presumptive taxation, a highly simplified variant of the Italian *analisi di settore* (see Arachi and Santoro 2007 for a description).

5. Data

Throughout this paper we rely on the Wage Survey (WS) of the National Employment Service. The WS is a linked employer-employee data set recently comprising observations on over 200,000 individuals in about 20,000 firms and budget institutions. The survey has been carried out annually since 1992. In the enterprise sector, the part of the survey that we use, the WS covers businesses employing at least 5 workers. All Hungarian firms employing more than 20 workers are obliged to report data for the WS while smaller firms are randomly selected from the census of enterprises. Firms employing 5-50 workers have to report individual data on each employee while large ones report data on a (roughly 10 per cent) random sample of their workers. The survey contains information on the wages and demographic and human capital variables of the workers and their job characteristics. In addition, financial variables (sales, assets, profits, and so on) are available on the firm level. The observations are weighted by the Employment Service to correct for the selection of firms and individuals.

Firms in the WS can be linked and followed over time. Individuals can not be linked directly but they can be identified across waves with acceptable precision using data on their firm ID, location of their workplace, year of birth, gender, education and 4-digit occupational code. We use several cross-sections of the WS and panels of individual and firm-level observations. A list is given below. Sample selection issues are dealt with in section 6.

Model	Data source	Number of observations
Preliminary transformation of 2003 wages (see section 4.2.)	Quasi-panel of individuals observed in WS 2000 and WS 2002, private sector	
DH estimates for 2003	WS 2003, private sector	100,809 workers
DH estimates for 2006	WS 2006, private sector	131,049 workers
Test 1	Panel of small firms observed in WS 2000 and WS 2003	263 firms
Residual wages for Test 1	WS 2000 and WS 2003	126,739 and 149,395 workers
Test 2	Quasi-panel of individuals observed in WS 2006 and WS 2007	7042 MW earners (as of 2006)
	WS firm panel 2006-2007	4150 firms

6. Results

6.1. Results of the DH estimates

If the DH model is correctly specified (including the distributional assumptions), then identification can be carried out even if X=Z, i.e. based merely on nonlinearities. However, to make the results more robust to deviations from the distributional assumptions, it is worth including variables which only influence the selection equation but not the wage equation. Therefore, in the selection equation we use occupational variables to capture grey employment directly (managerial and freelance- occupations, occupations with frequent cash transactions or jobs in trade, hotels, restaurants; see Table A1 in Appendix for definitions). Tables 2 and 3 present the DH estimates for 2003 and 2006.

Table 2, Table 3

Using the estimated parameters, the probability of under-reporting among MW earners and their genuine wages were calculated. The results suggest that in 2003 approximately half of the apparent MW earners hid part of their earnings from the tax authority. We estimate that the average „genuine" wage of the MW earners amounted to approximately 170 per cent of the MW and the average wage of cheating MW employees (using w>MW as the criterion for cheating) was around 220 per cent of MW. Similar results are obtained for 2006.

Table 4

It is interesting to analyze how employee and firm characteristics are related to under-reporting. Looking at *occupations*, under-reporting is more frequent where cash transactions are frequent, and in trade, hotels and restaurants (see the selection equation in Table 2). This is in line with expectations. Table 4 displays the estimated probabilities of under-reporting among MW earners, their average genuine wage by occupation and a "cheating indicator", which is the share of cheating MW earners among *all* employees. Using the definition w>MW for cheating, the estimated fraction of

cheaters among MW earners is only 18 per cent for cleaners, while it is around 65 per cent for blue collars in construction and approaches 100 per cent for managers and professionals. It is also clear that the share of MW earners is not a good indicator of under-reporting because fraud is frequent for some occupations with a high share of MW earners (e.g. in construction), while infrequent for others (cleaners, unskilled laborers).

Table 5

As far as *firm characteristics* are concerned, Table 5 displays the relation of economic sector, firm size and ownership to under-reporting. The "cheating indicator" (the share of cheating MW earners among all employees) is far the highest in construction and trade. Both the ratio of MW earners and cheating among MW earners are negatively correlated with firm size and with the share of foreign ownership (this can be seen from the selection equation in Table 2 as well).

The proportion of cheating *firms* (i.e. firms with at least one cheating employee) amounted to 38 per cent among firms employing 5-20 workers in 2003 (the sample for Test 1), and 25 per cent among all firms in 2006 (sample for Test 2).

6.2. Empirical specification and results of Test 1

We follow 263 small firms (5-20 employees) observed in the Wage Survey in 2000 as well as 2003.[12] Selection of firms to the panel is analyzed using the entire base-period sample. As shown by a probit estimate presented in Appendix Table A2, selection to the estimation sample was independent of the number of employees, level of wages, share of low-wage workers, sales revenues, profits, the capital-labor ratio, skill composition and location. On the basis of these estimates we regard the sample as randomly selected. Descriptive statistics are presented in Appendix Tables A3 and A4.

Tables 6-9

[12] We chose small firms because they report data on all of their employees, therefore their exposure to the MW hike and and their skill composition are precisely measured.

The estimates of equation (7) and its counterparts are presented in Tables 6-9. In Table 6 the observed rates of real average wage growth between 2000 and 2003 are regressed on the MW shock indicators interacted with the cheating proxies. The controls (relating to 2000) include the number of employees, profit/worker, a dummy for sales revenues falling short of material costs, fixed assets/worker, skill shares (primary, vocational, college), share of men, average age of the workforce, local unemployment rate and one-digit industry dummies. The table contains results from 48 regressions differing in the definitions of exposure to the MW hike, proxies of cheating behavior, inclusion/exclusion of controls and the samples considered (all firms versus only low-wage firms with at least one employee affected by the first MW hike). Consistent with the expectation, the point estimates are lower for cheating firms in all specifications. Reported wages grew faster than average in firms exposed to the MW hikes but they grew significantly faster in non-cheating than cheating firms. This is shown by F-tests for the equality of the parameters for cheaters and non-cheaters. [13]

Table 7 repeats the estimation using residual wages on the left hand i.e. wage growth controlled for changes in the composition of the workforce by education, experience and gender. The point estimates are always higher for non-cheating firms and the parameters for cheaters and non-cheaters differ significantly in the vast majority of the cases. When controls are included the parameter estimates prove to be significantly different in all model variants.

Table 8 looks at the effect of exposure to the MW hike on prospective employment growth. The point estimates for non-cheating firms are always negative and significant except in the last specification (when exposure is measured with the share of workers earning less than Ft 50,000 in May 2000). The estimates for cheating firms are always lower in absolute value and insignificant in the vast majority of the cases. The

[13] The distribution of individual wages within the two groups of firms, in 2003, provides further evidence that the DH model has predictive power. In non-cheating firms we find many workers paid slightly above the MW (hinting at spillover) while the cheaters paid almost half of their employees at the MW with only a few workers earning slightly more than that. See Appendix Figure A1.

coefficients for cheaters and non-cheaters can be considered statistically different in 28 out of the 48 specifications.[14]

Table 9 presents the estimates for changes in the share of unskilled workers (those with primary education or less) measured in percentage points. For non-cheating firms the effect of the MW hike proves to be negative and highly significant in all specifications. The point estimates for cheating firms are always smaller and insignificant in the majority of the model variants. The coefficients for cheaters and non-cheaters differ significantly in each of the 48 specifications.[15]

Generally, we find that firms strongly exposed to the minimum wage hike were more likely to raise wages fast and replace unskilled workers by skilled ones. Minimum wage effects appear to be highly significant and strong in the group of non-cheaters while the coefficients fall close to zero and/or insignificant in the group of suspected cheaters. The results for total employment are less robust but follow the same pattern.[16]

Differences in the size of the effects are difficult to summarize given the large number of the tested specifications. To illustrate the magnitudes we pick one specification (MW shock implied by the first hike, cheating suspected when w>MW, only low-wage firms, controls included). The coefficient for cheating firms in the wage change equation is close to unity (0.9492). Such an outcome is likely if firms raise the reported wages of the affected workers to Ft 40,000 and nothing else happens to reported wages, that is, if they continue or start to engage in under-reporting. By contrast, the coefficient for non-cheating firms is high above unity (1.5662) hinting at spillover effects. In the employment equation the coefficients are -0.4414 versus -1.1457, consistent with the

[14] We also estimated labor demand equations group by group following Machin et al (2003) i.e. by instrumenting actual wage growth with the MW shock variable. The output elasticities are identical for cheaters and non-cheaters (0.37/0.38) while the wage elasticity is -0.35 (significant at the 0.013 level) for non-chaters and insignificant for cheaters. This is consistent with the expectation that observed wage growth is a good predictor of actual wage growth in the former group but not in the latter. The results for the sample of low-wage firms are similar. We applied the classification based on w>MW.

[15] The results remain if we exclude firms employing no unskilled workers in the base period. Available on request.

[16] We checked how sales revenues and the capital-labor ratio were affected. Sales revenues are largely unaffacted by exposure to the MW hike. The capital-labor ratio is positively affected by exposure to the MW hike, the point estimates are higher for non-cheating firms but the estimates are statistically different in only less than half of the specifications.

expectation that cheating firms could contain labor costs and maintain more jobs at the cost of tax evasion. In Table 9 we find that in cheating firms exposure to the MW hike had no effect on the share of unskilled workers while in the group of non-cheaters one standard deviation difference in the MW shock (0.15) implied 1/3 standard deviation difference in the change of the unskilled share.

6.3. Responses to the introduction of a minimum contribution base

In Test 2 our starting point is the sample of MW earners in WS 2006 and observed in WS 2007. For the selection see Appendix Table A5. We estimate linear regressions for the change of the wage in 2006-2007 and probits (with the dependent variable coded 1 if the worker earned 2MW in 2007, and 0 otherwise). Our point of interest is the marginal effect of the 'cheater' dummies defined beforehand.

The results presented in Table 10 and 11 partly confirm our expectations. In Table 10 we look at how suspected cheating behavior affected wage change among the MW earners of 2006. We find highly significant effects in all specifications suggesting that cheating employers raised the wages of their MW earners significantly faster than non-cheating firms, and this result is robust to changes in the definition of cheating and the inclusion of controls. Base-period MW earners classified as cheaters (victims of cheating) were also more likely to earn 2MW in 2007 than their non-cheating counterparts but - when firm size and region dummies are included as controls - the coefficients are not significant at conventional levels in two equations out of three (Table 11).

Tables 10 and 11

Raising the reported wages of grey employees incurred costs and was expected to adversely affect output and employment. We estimate regressions to check if it was indeed the case. We find (Table 12) that the firm-level cheating proxies had a positive effect on (observed) average wages and significant negative effect on sales revenues and

employment. These results are robust to the changes of cheating proxies and inclusion of controls. [17]

Table 12

7. Conclusions

The DH model makes strong assumptions about the wage distribution, and finding variables, which affect selection to cheating without affecting wages, is also rather difficult. Reluctance to take the DH estimates at face value is therefore highly justified.

Our experiments aimed at testing if the DH estimates have predictive power. It seems that they worked well in the quasi-experimental settings analyzed in the paper: firms and workers suspected of tax evasion on the basis of the DH results responded differently to the exceptionally strong shocks under investigation. We obviously make both type 1 and type 2 errors in disentangling cheaters from non-cheaters but the results are encouraging for the analysis of 'grey employment', an important part of the informal economy in emerging market economies.

We believe that the results have practical importance. On the one hand, audits may be targeted by statistical profiles derived from the DH model, thereby improving compliance. However, by showing the loci of under-reporting the DH estimates also draw attention to the limits of tax enforcement. Disguised minimum wages have high shares in services provided to households and small businesses, freelance occupations, and small firm management – an attribute that limits the potential budgetary intakes from more stringent inspection. Cash transactions between households and the providers of personal services are difficult, if not impossible, to detect. Grey transactions of this kind can rather be whitened indirectly, by creating incentives to require receipts and making clear the link between reported income and access to publicly financed services and transfers such as pensions.

[17] We found the effects on the capital-labor ratio to be weaker and (in several specifications) insignificant in the short run covered by the data.

On the other hand, the DH results call for more cautious MW policies. The micro-data do not support the popular belief that in Hungary 'millions' are fraudulently paid the minimum wage – an assumption that served as a justification for regulations like a minimum contribution to be paid after 2MW. Reducing the under-reporting of wages by means of substantially increasing the MW and/or the tax burden on it is an undoubtedly cheap alternative to independent checks. However, raising the costs of low-wage employment across the board is a poorly targeted policy, which can further reduce unskilled job opportunities: an undesirable outcome in a country, where 6 out of 10 low-educated prime-age adults are out of work.

References

Allingham, M. and A. Sandmo (1972): Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1, 323-338.

Arachi, G. and Santoro, A. (2007): Tax enforcement for SMEs. Lessons from the Italian experience? *eJournal of Tax Research* 5(2), 225–243.

Benedek D., Rigó M., Scharle Á. and Szabó P. (2006): Increases in the minimum wage in Hungary, 2001-2006, *Working Papers in Public Finance* No. 16.

Cragg, J. (1971): Some statistical models for limited dependent variables with application to the demand for durable goods, *Econometrica* 39, 829-844.

Dickens, R., Machin, S. and Manning, A. (1994): The effects of minimum wages on employment: theory and evidence from the UK, *NBER Working Paper* 4742, Cambridge, MA.

Erdogdu, S. (2009): Turkey – Minimum wage in tension between economic and social concerns, in: Daniel Vaughan-Whitehead (ed.): The minimum wage in the enlarged EU, Edward Elgar 2009

European Commission (2007): Undeclared work in the European Union. Report, Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions.

Halpern, L., M. Koren, G. Kőrösi and J. Vincze (2004): A minimálbér költségvetési hatásai. (The budgetary effects of the minimum wage, in Hungarian), *Közgazdasági Szemle* 51(4), 325–345.

Kertesi G. and J. Köllő (2003): Fighting low equilibria by doubling the minimum wage – Hungary's experiment, *IZA Discussion Paper* No. 970, Bonn

Kriz K. A., Meriküll J., Paulus, A. and Staehr, K. (2007): Why do Individuals Evade Payroll and Income Taxation in Estonia? *U. of Tartu Economics and Business Administration Working Paper* No. 49-2007.

Labeaga, J. M. (1999): A double-hurdle rational addiction model with heterogeneity: Estimating the demand for tobacco, *Journal of Econometrics* 93, 49-72.

Machin, S., Manning, A. and Rahman, L. (2003): Where the minimum wage bites hard: Introduction of minimum wages in a low wage sector, *Journal of the European Economic Association* 1(1), 154–180.

Madzharova, B. (2010): The effect of low corporate income tax on social security evasion, Cerge-EI Prague, EPRN Seminars, Univ. of Copenhagen, 2010 Nov 12, http://web.econ.ku.dk/eprn_epru/Seminar/EPRU%20Nov12.Madzharova.pdf

Martinez-Espineira, R. (2006): A Box-Cox double hurdle model of wildlife valuation, the citizen's perspective, *Ecological Economics* 58, 192-208.

Masso, J. and Krillo, K. (2009): Estonia, Latvia and Lithuania – Minimum wages in the context of migration and labour shortages, in: Daniel Vaughan-Whitehead (ed.): The minimum wage in the enlarged EU, Edward Elgar 2009.

Meriküll J. and Staehr. K (2010): Unreported Employment and Envelope Wages in Mid-Transition: Comparing Developments and Causes in the Baltic Countries, *Comparative Economic Studies* 52 (September), 637–670.

Meyer, R. H. and Wise, D. A. (1983a): Discontinuous distributions and missing persons: the minimum wage and unemployed youth, *Econometrica* 51, 1677-1698.

Meyer, R. H. and Wise, D. A. (1983b): The effects of minimum wage on employment and earnings of youth, *Journal of Labor Economics* 1, 66-100.

Moffatt, P. G. (2005): Hurdle models of loan default, *Journal of the Operational Research Society* 56, 1063-1071.

Saz-Salazar, S. D. and Rausell-Köster, P. (2006): A double-hurdle model of urban green areas valuation: Dealing with zero responses, *Landscape and Urban Planning* 84, 241–251.

Shelkova, N. Y. (2008): Low wage labor markets and the power of suggestion, Working Paper 2008/33, *Department of Economics, University of Connecticut.*

Szabó, P. A. (2007): A 2000–2001. évi minimálbér-emelés hatása a jövedelemeloszlásra (The effect of the 2000-2001 minimum wage increase on income distribution, in Hungarian), *Közgazdasági Szemle* 54 (May), 397–414.

Teklewold, H., Dadi, L., Yami, A. and Dana, N. (2006): Determinants of adoption of poultry technology: a double-hurdle approach, *Livestock Research for Rural Development* 18.

Tonin, M. (2006): The effects of the minimum wage in an economy with tax evasion, *IIES, Stockholm University*, August, 2006.

Tonin, M. (2007): Minimum wage and tax evasion: theory and evidence, Institute of Economics, Budapest, DP 2007/1.

World Bank (2005): Enhancing Job Opportunities: Eastern Europe and the Former Soviet Union, *World Bank*, Washington.

Yaniv, G., 1988. Withholding and non-withheld tax evasion. *Journal of Public Economics* 35, 183-204.

Yen, S. T. and Jones, A. M. (1997): Household consumption of cheese, an inverse hyperbolic sine double hurdle approach with dependent errors, *American Journal of Agricultural Economics* 79, 246-251.

Table 1. Fraction paid the exact amount of the minimum wage in May 2003

	Per cent paid the MW	Composition *All MW earners=100*
Top managers	11.9	1.6
Managers (heads of department, foremen, etc.)	4.7	3.2
Managers of small firms	27.4	3.4
Engineers	3.4	1.0
Architects and construction technicians	17.9	0.5
Professionals in health, education and social services (private)	2.1	0.1
Other professionals	3.6	0.6
Lawyers, business and tax advisors, accountants	14.9	1.5
Freelance cultural occupations (musicians, actors, writers, artists, etc.)	15.8	0.5
Technicians	6.8	2.8
Administrators	5.0	5.2
Agents, brokers	17.7	1.2
Office workers	9.0	4.9
Blue collars in trade and catering	25.2	16.3
Blue collars in transport	3.0	0.5
Services A (other than B and C)	7.2	1.1
Services B (health and social services, private)	0.4	0.0
Services C (personal services)	22.0	1.0
Farmers and farm workers	10.4	3.4
Blue collars in heavy industry and engineering	5.4	6.9
Blue collars in light industry	9.5	10.3
Blue collars in construction (structural construction, house building)	20.6	10.8
Blue collars in construction (civil engineering)	4.3	0.2
Assemblers, machine operators	2.6	4.0
Truck drivers	17.5	4.6
Porters, guards, cleaners	13.6	8.4
Unskilled laborers, casual workers	23.1	6.3
Total	9.4	100.0

Source: Wage Survey, 2003. Number of observations = 201,971. For this table several occupations were divided into parts on the basis of industrial affiliation and firm size in order to capture differences in the scope for cash transactions with customers (house building versus civil engineering, personal versus other kind of services, small firm versus large firm managers, and so on).

Table 2: OLS and DH estimates of wages in 2003

	OLS		Double hurdle	
	Coefficient	St. error	Coefficient	St. Error
Wage equation				
Ezperience	0.0170***	0.0006	0.0191***	0.0006
Exp squared	-0.0002***	0.0000	-0.0002***	0.0000
Male	0.1328***	0.0037	0.1526***	0.0038
Vocational edu.	0.0551***	0.0049	0.0571***	0.0049
Secondary edu.	0.1663***	0.0056	0.1821***	0.0057
Higher edu.	0.5290***	0.0081	0.5636***	0.0082
Agriculture	-0.5175***	0.0292	-0.4708***	0.0286
Construction	-0.6761***	0.0282	-0.5812***	0.0278
Services	-0.6507***	0.0290	-0.5990***	0.0282
Trade	-0.6996***	0.0276	-0.6632***	0.0270
Industry	-0.6013***	0.0276	-0.5499***	0.0268
Cleaners	-0.8314***	0.0291	-0.8216***	0.0284
Unskilled laborers	-0.7577***	0.0283	-0.7423***	0.0276
Machine operators	-0.6071***	0.0278	-0.5517***	0.0270
Porters and guards	-0.8028***	0.0296	-0.7766***	0.0290
Drivers	-0.6256***	0.0283	-0.5234***	0.0276
Office clerks	-0.5235***	0.0278	-0.4578***	0.0271
Technicians, assistants	-0.3993***	0.0277	-0.3406***	0.0268
Administrators	-0.4570***	0.0273	-0.3823***	0.0265
Managers	-0.0957***	0.0271	-0.0052	0.0263
Professionals	-0.3204***	0.0274	-0.2670***	0.0265
Architects and construction technicians	-0.6305***	0.0374	-0.4616***	0.0380
Budapest	0.0970***	0.0036	0.1351***	0.0038
Value added per worker (log)	0.1915***	0.0018	0.1907***	0.0018
Fixed assets per worker (log)	0.0101***	0.0012	0.0135***	0.0012
Firm of foreign ownership	0.2053***	0.0040	0.1860***	0.0039
Firm with 5-10 employees	-0.5114***	0.0059	-0.3664***	0.0078
Firm with 11-20 employees	-0.4232***	0.0052	-0.3163***	0.0062
Firm with 21-50 employees	-0.2488***	0.0042	-0.1891***	0.0044
Firm with 51-300 employees	-0.1155***	0.0042	-0.0971***	0.0041
Inverse Mill's ratio*	-0.0492***	0.0066	-0.0492***	0.0068
Constant	0.9627***	0.0295	0.8522***	0.0288
Selection equation				
Firm of foreign ownership	-	-	0.6619***	0.0366
Managerial and freelance**	-	-	-0.7897***	0.0541
Cash transactions ***	-	-	-0.5977***	0.0234
Retail trade	-	-	-0.3025***	0.0332
Budapest	-	-	-0.3294***	0.0263
Works in a city	-	-	0.0581**	0.0284
Works in a village	-	-	0.0476	0.0404
Number of firms (log)	-	-	-0.7180***	0.0530
Tourism (nights per inhabitant) in the municipality	-	-	0.0276***	0.0037

Income tax base per capita in the municipality	-	-	0.9456***	0.0599
Firm with 5-10 employees			-2.2573***	0.0957
Firm with 11-20 employees			-2.0227***	0.0953
Firm with 21-50 employees			-1.6207***	0.0949
Firm with 51-300 employees			-1.0373***	0.0972
Constant	-	-	0.4949	0.3074
N of observations	100 809		100 809	
Adjusted R^2	0.5556		-	
Rho	-		-0.1727	

*** p<0.01, ** p<0.05, * p<0.1.

* Because of the negative employment effects of the high MW some people are crowded out of employment. To tackle this sample selection issue, we run a probit model on the Labour Force Survey to estimate the probability of being employed (being in the LEED sample), and use the inverse Mill's ratio from the probit regression in the wage equation of the DH model.

** Dummy variable for managerial and freelance occupations. Freelance occupations include professionals in culture and art, agents and brokers

***Dummy variable for occupations, where cash transactions are assumed to occur frequently. Includes car mechanics, electricians, plumbers, household employees, couriers, truck drivers and workers in personal services and house building.

Table 3: OLS and DH estimates of wages in 2006

	OLS		Double hurdle	
	Coefficient	St. error	Coefficient	St. Error
Wage equation				
Ezperience	0.0193***	0.000428	0.0210***	0.000459
Exp squared	-0.000291***	8.71e-06	-0.000321***	9.35e-06
Male	0.131***	0.00299	0.151***	0.00317
Vocational edu.	0.0545***	0.00406	0.0638***	0.00433
Secondary edu.	0.149***	0.00449	0.166***	0.00480
Higher edu.	0.494***	0.00648	0.523***	0.00690
Agriculture	-0.383***	0.0201	-0.428***	0.0211
Construction	-0.352***	0.0189	-0.376***	0.0199
Services	-0.344***	0.0192	-0.399***	0.0200
Trade	-0.423***	0.0181	-0.482***	0.0190
Industry	-0.307***	0.0181	-0.352***	0.0188
Cleaners	-0.536***	0.0192	-0.631***	0.0201
Unskilled laborers	-0.501***	0.0188	-0.603***	0.0196
Machine operators	-0.298***	0.0183	-0.335***	0.0191
Porters and guards	-0.535***	0.0204	-0.611***	0.0213
Drivers	-0.293***	0.0188	-0.317***	0.0197
Office clerks	-0.253***	0.0184	-0.276***	0.0192
Technicians, assistants	-0.103***	0.0182	-0.135***	0.0189
Administrators	-0.172***	0.0178	-0.194***	0.0185
Managers	0.161***	0.0176	0.151***	0.0183
Professionals	0.0239	0.0179	-0.0212	0.0185
Architects and construction	-0.0578**	0.0293	-0.0985***	0.0299

technicians				
Budapest	0.0873***	0.00289	0.103***	0.00334
Value added per worker log	0.163***	0.00150	0.174***	0.00159
Fixed assets per worker log	0.0127***	0.000827	0.00897***	0.000885
Firm of foreign ownership	0.270***	0.00339	0.248***	0.00359
Firm with 5-10 employees	-0.405***	0.00606	-0.343***	0.00868
Firm with 11-20 employees	-0.332***	0.00506	-0.300***	0.00665
Firm with 21-50 employees	-0.214***	0.00377	-0.192***	0.00496
Firm with 51-300 employees	-0.156***	0.00315	-0.143***	0.00419
Constant	0.532***	0.0190	0.539***	0.0199
Selection equation				
Firm of foreign ownership	-	-	0.826***	0.0354
Managerial and freelance*	-	-	-0.646***	0.0505
Cash transactions*	-	-	-0.488***	0.0228
Retail trade	-	-	-0.213***	0.0359
Budapest	-	-	-0,3294***	0,0263
Works in a city	-	-	0.174***	0.0272
Works in a village	-	-	-0.0306	0.0339
Tourism nights per inhabitant in the municipality	-	-	-0.00226***	0.000755
Income tax base per capita in the municipality	-	-	0.320***	0.0453
Firm with 5-10 employees			-2.098***	0.104
Firm with 11-20 employees			-1.745***	0.103
Firm with 21-50 employees			-1.560***	0.102
Firm with 51-300 employees			-1.419***	0.101
Constant	-	-	1.055***	0.318
N of observations	100 809		132,933	
Adjusted R^2	0.5556		-	
Rho	-		-0.1940	

*** p<0.01, ** p<0.05, * p<0.1.

*) See the notes to Table 2

Table 4: Estimated under-reporting among MW earners in different occupations

	Estimated probability of under-reporting among MW earners (per cent)			Share of MW earners (per cent)	Cheating indicator			Simulated wage of cheaters		
Occupations	w>MW	w>1.5*MW	P>0.5		w>MW	w>1.5*MW	P>0.5	w>MW	w>1.5*MW	P>0.5
Agriculture	47.2	25.8	41.7	14.4	6.8	3.7	6.0	89,081	112, 924	79,016
Construction	66.4	36.8	82.8	27.3	18.1	10.1	22.6	90,243	113,583	80,064
Services	51.3	31.4	51.9	8.5	4.4	2.7	4.4	104,622	133,137	94,129
Trade	42.7	18.5	34.0	25.8	11.0	4.8	8.8	80,884	109,118	75,679
Industry	47.6	28.9	43.8	15.6	7.4	4.5	6.8	97,017	120,324	87,679
Other blue-collar										
Cleaners	18.3	6.6	5.8	24.6	4.5	1.6	1.4	73,755	97,902	71,092
Unskilled laborers	29.0	13.1	11.2	29.3	8.5	3.9	3.3	80,929	105,501	75,897
Machine operators	43.4	25.0	39.0	7.6	3.3	1.9	3.0	90,742	112,403	78,855
Porters and guards	36.4	18.9	19.9	16.3	5.9	3.1	3.2	92,629	121,527	88,413
Drivers	77.5	51.5	91.3	17.9	13.9	9.2	16.3	101,358	121,292	92,039
White-collar										
Office clerks	58.5	40.5	63.7	13.5	7.9	5.5	8.6	109,363	130,582	96,593
Technicians, assistants	79.2	63.8	88.8	6.6	5.2	4.2	5.9	129,683	145,855	120,246
Administrators	75.5	60.6	89.2	8.2	6.2	5.0	7.3	129,023	145,138	115,650
Managers	95.4	90.0	98.8	7.4	7.1	6.7	7.3	197,103	205,113	191,458
Professionals	98.5	93.5	100.0	3.9	3.8	3.6	3.9	189,656	196,317	187,571
Architects and construction technicians	93.9	81.6	98.0	16.1	15.1	13.1	15.8	181,875	199,437	176,380
Teachers and doctors	100.0	100.0	100.0	1.7	1.7	1.7	1.7	224,079	224,079	224,079
Total	**54.5**	**35.4**	**53.7**	**14.1**	**7.7**	**5.0**	**7.6**	**112,303**	**139,995**	**105,850**

Source: Wage Survey, 2003.
Cheating indicator: share of cheating MW earners among all employees
Pro memo: the MW was Ft 50,000 in 2003

Table 5: Estimated under-reporting among MW earners by firm characteristics

	Estimated probability of under-reporting among MW earners (per cent)			Share of MW earners (per cent)	Cheating indicator			Simulated wage of cheaters		
Firm characteristics	w>MW	w>1.5*MW	P>0.5		w>MW	w>1.5*MW	P>0.5	w>MW	w>1.5*MW	P>0.5
Firm size										
5-10 employees	63.5	38.1	67.8	37.5	23.8	14.3	25.4	105 621	135 575	98 747
11-20 employees	59.2	36.1	60.8	28.2	16.7	10.2	17.1	106 286	135 495	99 563
21-50 employees	53.2	36.4	49.1	17.2	9.1	6.3	8.4	115 174	139 939	110 269
51-300 employees	40.2	32.8	32.7	6.7	2.7	2.2	2.2	141 745	159 904	143 960
301- employees	6.5	5.9	3.2	1.6	0.1	0.1	0.1	211 244	224 928	220 781
Economic branch										
Agricult. and fishing	47.4	28.6	41.2	11.6	5.5	3.3	4.8	107 440	137 773	103 424
Mining	61.3	41.9	64.5	7.6	4.7	3.2	4.9	122 393	151 889	113 436
Manufact.	48.2	31.2	44.8	11.7	5.6	3.6	5.2	109 927	136 855	103 455
Electricity gas wand water supply	33.3	15.4	30.8	1.0	0.3	0.2	0.3	93 673	125 599	74 867
Construction	58.2	36.5	62.9	27.1	15.8	9.9	17.1	107 265	134 662	97 662
Trade	57.4	35.7	56.7	22.3	12.8	8.0	12.7	108 051	136 666	102 187
Hotels and restaurants	43.9	24.7	34.1	25.9	11.4	6.4	8.8	100 638	133 572	102 734
Transport	74.8	54.7	84.4	6.8	5.1	3.7	5.7	119 426	140 313	109 000
Financial intermed.	81.8	72.7	90.9	1.9	1.6	1.4	1.7	165 875	178 137	154 288
Real estate, business activities	59.8	46.0	60.1	14.0	8.4	6.4	8.4	138 871	161 999	132 982
Other community services	42.6	29.1	37.7	8.3	3.5	2.4	3.1	118 736	146 148	118 547
Ownership										
Domestic	56.1	36.1	55.5	19.8	11.1	7.1	11.0	110 024	137 139	103 375
Foreign	44.9	39.1	45.3	2.8	1.3	1.1	1.3	177 596	194 756	170 560
Mixed	49.3	30.2	46.9	18.6	9.2	5.6	8.7	110 439	141 929	105 722
Total	**54.5**	**35.4**	**53.7**	**14.1**	**7.7**	**5.0**	**7.6**	**112303**	**139995**	**105850**

Source: Wage Survey, 2003.
Cheating indicator: share of cheating MW earners among all employees
Pro memo: the MW was Ft 50,000 in 2003

Table 6: Exposure to the MW hikes and prospective average wage growth

Dependent variable: log change of the real average wage in 2000-2003. OLS.

Proxy of cheating	Controls	β_1 (cheating firms)	β_2 (non-cheating firms)	F-test H_0: $\beta_1=\beta_2$	β_1 (cheating firms)	β_2 (non-cheating firms)	F-test H_0: $\beta_1=\beta_2$
Sample:			All firms (N=263)			Low-wage firms[a] (N=194)	
Proxy of exposure: Average wage increase 'implied' by the first MW hike[b]							
P>0.5	No	0.9567***	1.5137***	10.35***	0.8694***	1.4211***	10.19***
	Yes	0.9365***	1.5048***	10.29***	0.9665***	1.5171***	9.86***
w>MW	No	0.9728***	1.5412***	9.14***	0.8844***	1.4483***	9.04***
	Yes	0.9352***	1.5413***	10.89***	0.9492***	1.5662***	11.83***
w>1.5MW	No	1.0164***	1.4074***	5.76**	0.9304***	1.3139***	5.52**
	Yes	0.9440***	1.4321***	8.95***	0.9441***	1.4762***	11.33***
Proxy of exposure: Average wage increase 'implied' by the first and the second MW hikes[b]							
P>0.5	No	0.6627****	1.0115***	12.50***	0.6509***	0.9966***	12.14***
	Yes	0.6892***	1.0519***	12.91***	0.7033***	1.0506***	11.54***
w>MW	No	0.6722***	1.0276***	11.22***	0.6595***	1.0112***	10.92***
	Yes	0.6898***	1.0666***	13.39***	0.6939***	1.0785***	14.03***
w>1.5MW	No	0.7000***	0.9399***	6.59**	0.6883***	0.9241***	6.27**
	Yes	0.6949***	0.9944***	10.44***	0.6931***	1.0141***	12.32***
Proxy of exposure: Fraction of employees earning less than Ft 40,000 in 2000							
P>0.5	No	0.3838***	0.5640***	10.74***	0.4062***	0.5879***	10.75***
	Yes	0.3759***	0.5576***	10.70***	0.4424***	0.6159***	6.68***
w>MW	No	0.3889***	0.5719***	9.81***	0.4109***	0.5951***	9.82***
	Yes	0.3760***	0.5649***	11.24***	0.4359***	0.6272***	11.49***
w>1.5MW	No	0.4163***	0.5156***	3.36*	0.4366***	0.5374***	3.40*
	Yes	0.3854***	0.5192***	6.35**	0.4371***	0.5855***	8.17***
Proxy of exposure: Fraction of employees earning less than Ft 50,000 in 2000							
P>0.5	No	0.3870***	0.5375***	10.74***	0.4701***	0.6251***	11.00***
	Yes	0.3922***	0.5494***	11.26***	0.4524***	0.6085***	10.60**
w>MW	No	0.3899***	0.5457***	10.44***	0.4722***	0.6324***	10.64***
	Yes	0.3921***	0.5526***	12.01***	0.4472***	0.6164***	13.24***
w>1.5MW	No	0.4098***	0.5008***	3.89**	0.4922***	0.5839***	3.87*
	Yes	0.3967***	0.5189***	7.27***	0.4472***	0.5842***	9.17***

Significant at the *) 0.1 **) 0.05 ***) 0.01 level

a) At least one worker paid less than Ft 40,000 (50,000) in May 2000

b) Log average wage increase (ω) in the hypothetical case of paying Ft 40,000 (50,000) to workers affected by the MW hikes and leaving other wages unchanged. Pro memo: the MW increased from Ft 25,500 to Ft 40,000 in January 2001 and Ft 50,000 in January 2002.

Controls (all variables relate to 2000): number of employees, profit/worker, dummy for sales revenues falling short of material costs, fixed assets/worker, skill shares (primary, vocational, college), share of men, average age, local unemployment rate, one-digit industry dummies

Table 7: Exposure to the MW hikes and prospective residual wage[a] growth

Dependent variable: log change of the residual average wage in 2000-2003. OLS.

Proxy of cheating	Controls	β_1 (cheating firms)	β_2 (non-cheating firms)	F-test H_0: $\beta_1=\beta_2$	β_1 (cheating firms)	β_2 (non-cheating firms)	F-test H_0: $\beta_1=\beta_2$
Sample:			All firms (N=263)			Low-wage firms[b] (N=194)	
Proxy of exposure: Average wage increase 'implied' by the first MW hike[c]							
P>0.5	No	0.8024***	1.0903***	2.83*	0.7003***	0.9818***	2.76*
	Yes	0.8000***	1.1594***	4.33**	0.7571***	1.1218***	4.60**
w>MW	No	0.8130***	1.1015***	2.45	0.7099***	0.9930***	2.41
	Yes	0.8041***	1.1755***	4.44**	0.7527***	1.1444***	5.33**
w>1.5MW	No	0.8343***	1.0344***	1.56	0.7336***	0.9249***	6.93***
	Yes	0.8084***	1.1096***	3.54*	0.7483***	1.0884***	4.88**
Proxy of exposure: Average wage increase 'implied' by the first and the second MW hikes[c]							
P>0.5	No	0.5575***	0.7548***	4.13**	0.5245***	0.7162***	3.89**
	Yes	0.5974***	0.8438***	6.38**	0.5713***	0.8153***	6.18**
w>MW	No	0.5619***	0.7651***	3.85*	0.5283***	0.7261***	3.65*
	Yes	0.6005***	0-8499***	6.48**	0.5690***	0.8288***	7.19***
w>1.5MW	No	0.5792***	0.7141***	2.18	0.5461***	0.6752***	1.98
	Yes	0.6058***	0.8000***	4.64**	0.5716***	0.7822***	5.58**
Proxy of exposure: Fraction of employees earning less than Ft 40,000 in 2000							
P>0.5	No	0.3400***	0.4498***	4.54**	0.3543***	0.4650***	4.55**
	Yes	0.3512***	0.4846***	6.82***	0.3877***	0.5209***	6.78***
w>MW	No	0.3419***	0.4563***	4.45**	0.3559***	0.4712***	4.46**
	Yes	0.3527***	0.4883***	7.06***	0.3855***	0.5257***	7.81***
w>1.5MW	No	0.3609***	0.4194***	1.34	0.3739***	0.4332***	1.37
	Yes	0.3618***	0.4533***	3.58*	0.3894***	0.4925***	4.87**
Proxy of exposure: Fraction of employees earning less than Ft 50,000 in 2000							
P>0.5	No	0.3379***	0.4311***	4.68**	0.3972***	0.4892***	4.37**
	Yes	0.3674***	0.4820***	7.16***	0.3968***	0.5151***	7.27***
w>MW	No	0.3377***	0.4387***	5.08**	0.3964***	0.4962***	4.74**
	Yes	0.3690***	0.4821***	7.28***	0.3955***	0.5174***	8.39***
w>1.5MW	No	0.3540***	0.4065***	1.50	0.4131***	0.4622***	1.27
	Yes	0.3759***	0.4554***	3.77*	0.3999***	0.4903***	4.84**

Significant at the *) 0.1 **) 0.05 ***) 0.01 level

a) Residual wage = firm-level mean residuals from standard Mincer-equations (log wages regressed on gender, experience, expericence squared and years in school) using individual data of the 2000 and 2003 waves of the Wage Survey.

b) At least one worker paid less than Ft 40,000 (50,000) in May 2000

c) Log average wage increase (ω) in the hypothetical case of paying Ft 40,000 (50,000) to workers affected by the MW hikes and leaving other wages unchanged. Pro memo: the MW increased from Ft 25,500 to Ft 40,000 in January 2001 and Ft 50,000 in January 2002.

Controls (all variables relate to 2000): number of employees, profit/worker, dummy for sales revenues falling short of material costs, fixed assets/worker, skill shares (primary, vocational, college), share of men, average age, local unemployment rate, one-digit industry dummies

Table 8: Exposure to the MW hikes and prospective employment growth

Dependent variable: log change of the number of employees in 2000-2003. OLS.

Proxy of cheating	Controls	β_1 (cheating firms)	β_2 (non-cheating firms)	F-test H_0: $\beta_1=\beta_2$	β_1 (cheating firms)	β_2 (non-cheating firms)	F-test H_0: $\beta_1=\beta_2$
Sample:			All firms (N=263)			Low-wage firms[a] (N=194)	
Proxy of exposure: Average wage increase 'implied' by the first MW hike[b]							
P>0.5	No	-0.0586	-0.8496***	6.46**	-0.1780	-0.9765***	6.59**
	Yes	-0.1929	-0.8735**	4.45**	-0.4276*	-1.1279***	4.75**
w>MW	No	-0.0913	-0.8751***	5.75**	-0.2105	-1.0004***	5.87**
	Yes	-0.1971	-0.9111**	4.41**	-0.4414*	-1.1457***	4.27**
w>1.5MW	No	-0.1917	-0.6497**	2.22	-0.4783**	-0.9888***	2.62
	Yes	-0.2307	-0.7452**	2.68	-0.3084	-0.7768**	2.32
Proxy of exposure: Average wage increase 'implied' by the first and the second MW hikes[b]							
P>0.5	No	-0.0096	-0.4558**	5.85**	-0.0308	-0.4786**	5.87**
	Yes	-0.0938	-0.4621**	3.76*	-0.2877*	-0.6724***	4.04**
w>MW	No	-0.0332	-0.4605**	4.89**	-0.0539	-0.4827**	4.91**
	Yes	-0.0980	-0.4734**	3.55*	-0.2944*	-0.6808***	3.69*
w>1.5MW	No	-0.0846	-0.3383*	1.96	-0.1043	-0.3598*	1.98
	Yes	-0.1033	-0.3800*	2.41	-0.2949*	-0.6050**	2.77*
Proxy of exposure: Fraction of employees earning less than Ft 40,000 in 2000							
P>0.5	No	0.0044	-0.2308**	5.12**	-0.0665	-0.3061**	5.29**
	Yes	-0.0278	-0.2022*	2.77*	-0.1023	-0.2870**	3.07*
w>MW	No	-0.0106	-0.2299**	4.12**	-0.0807	-0.3042**	4.25**
	Yes	-0.0314	-0.2055*	2.49	.0.1087	-0.2821*	2.36
w>1.5MW	No	-0.0477	-0.1586*	1.17	-0.1154	-0.2308**	1.25
	Yes	-0.0333	-0.1628	1.54	-0.1097	-0.2383*	1.46
Proxy of exposure: Fraction of employees earning less than Ft 50,000 in 2000							
P>0.5	No	0.0295	-0.1538*	4.78**	0.0227	-0.1611	4.75**
	Yes	0.0489	-0.0744	2.11	0.0487	-0.0788	2.15
w>MW	No	0.0149	-0.1496*	3.62*	0.0092	-0.1558	3.59*
	Yes	0.0476	-0.0764	1.98	0.0487	-0.0776	1.92
w>1.5MW	No	-0.0048	-0.1031	1.38	-0.0085	-0.1071	1.38
	Yes	0.0583	-0.0564	1.81	0.0634	-0.0613	2.03

Significant at the *) 0.1 **) 0.05 ***) 0.01 level

a) At least one worker paid less than Ft 40,000 (50,000) in May 2000

b) Log average wage increase (ω) in the hypothetical case of paying Ft 40,000 (50,000) to workers affected by the MW hikes and leaving other wages unchanged. Pro memo: the MW increased from Ft 25,500 to Ft 40,000 in January 2001 and Ft 50,000 in January 2002.

Controls (all variables relate to 2000): profit/worker, dummy for sales revenues falling short of material costs, fixed assets/worker, skill shares (primary, vocational, college), share of men, average age, local unemployment rate, one-digit industry dummies

Table 9: Exposure to the MW hikes and prospective change in the share of unskilled workers[a]

Dependent variable: percentage points change in the share of unskilled workers 2000-2003. OLS.

Proxy of cheating	Controls	β_1 (cheating firms)	β_2 (non-cheating firms)	F-test H_0: $\beta_1=\beta_2$	β_1 (cheating firms)	β_2 (non-cheating firms)	F-test H_0: $\beta_1=\beta_2$
Sample:			All firms (N=263)			Low-wage firms[b] (N=194)	
Proxy of exposure: Average wage increase 'implied' by the first MW hike[c]							
P>0.5	No	-0.0339	-0.4612***	8.97***	-0.0449	-0.4727***	8.91***
	Yes	-0.1615**	-0.5623***	9.32***	-0.2472***	-0.6401***	9.32***
w>MW	No	-0.0544	-0.4709***	7.66***	-0.0649	-0.4719***	7.62***
	Yes	-0.1829**	-0.5613***	7.57***	-0.2748***	-0.6291***	7.06***
w>1.5MW	No	-0.0073	-0.4529***	10.88***	-0.0196	-0.4663***	10.80***
	Yes	-0.1460*	-0.5344***	10.27***	-0.2440***	-0.6105***	9.87***
Proxy of exposure: Average wage increase 'implied' by the first and the second MW hikes[c]							
P>0.5	No	-0.0314	-0.2604***	7.94***	-0.0493	-0.2776***	7.78***
	Yes	-0.1191**	-0.3412***	8.53***	-0.2102***	-0.4336***	8.83***
w>MW	No	-0.0438	-0.2625***	6.51**	-0.0610	-0.2789***	6.39**
	Yes	-0.1331**	-0.3354***	6.60**	-0.2280***	-0.4225***	6.41**
w>1.5MW	No	-0.0112	-0.2578***	10.26***	-0.0298	-0.2759***	10.06***
	Yes	-0.1063*	-0.3232***	9.70***	-0.2036***	-0.4147***	9.81***
Proxy of exposure: Fraction of employees earning less than Ft 40,000 in 2000							
P>0.5	No	-0.0305	-0.1372***	6.38***	-0.0479	-0.1558***	6.45**
	Yes	-0.0799**	-0.1865***	6.88***	-0.1404***	-0.2477***	6.98***
w>MW	No	-0.0366	-0.1379***	5.22**	-0.0546	-0.1559***	5.27**
	Yes	-0.0879***	-0.1828***	5.21**	-0.1506***	-0.2394***	4.87**
w>1.5MW	No	-0.0174	-0.1376***	8.85***	-0.0351	-0.1564***	8.93***
	Yes	-0.0717**	-0.1758***	7.69***	-0.1356***	-0.2316***	7.00***
Proxy of exposure: Fraction of employees earning less than Ft 50,000 in 2000							
P>0.5	No	-0.0225	-0.0934***	4.25**	-0.0439	-0.1162**	4.22**
	Yes	-0.0524*	-0.1260***	4.51**	-0.0995*	-0.1748***	4.57**
w>MW	No	-0.0268	-0.0936**	3.48*	-0.0478	-0.1158**	3.47*
	Yes	-0.0586*	-0.1227***	3.38*	-0.1067**	-0.1673***	3.06*
w>1.5MW	No	-0.0092	-0.0974***	7.20***	-0.0304	-0.1198***	7.10***
	Yes	-0.0426	-0.1217***	6.27**	-0.0918*	-0.1672***	5.89**

Significant at the *) 0.1 **) 0.05 ***) 0.01 level

a) Unskilled: workers with complete or incomplete primary school attainment

b) At least one worker paid less than Ft 40,000 (50,000) in May 2000

c) Log average wage increase (ω) in the hypothetical case of paying Ft 40,000 (50,000) to workers affected by the MW hikes and leaving other wages unchanged. Pro memo: the MW increased from Ft 25,500 to Ft 40,000 in January 2001 and Ft 50,000 in January 2002.

Controls (all variables relate to 2000): number of employees, profit/worker, dummy for sales revenues falling short of material costs, fixed assets/worker, share of men, average age, local unemployment rate, one-digit industry dummies

Table 10: The effect of estimated cheating behavior on wage growth between May 2006 and May 2007

(OLS, 7042 observations)

	Proxy used					
	P>0.5		w>MW		w>1.5 MW	
Controls	Coefficient	St. error	Coefficient	St. error	Coefficient	St. error
No	*19825****	1552	*8703****	1072	*10706****	1268
Education	15699***	1275	5742***	980.1	7263***	1141
All	12095***	1339	3472***	1014	5089***	1173

*** p<0.01, ** p<0.05, * p<0.1.

Controls (all variables relate to 2006): Dummies for education (college graduate, secondary school and vocational school), work experience in years, dummies for gender, municipality and firm size categories

Table 11: The effect of estimated cheating behavior on the probability that a worker paid the MW in May 2006 was paid 2MW in May 2007

(Probit marginal effects, 7042 observations)

	Proxy used					
	P>0.5		w>MW		w>1.5 MW	
Controlls	Coefficient	St. error	Coefficient	St. error	Coefficient	St. error
No	0.133***	0.0175	0.0479***	0.00903	0.0583***	0.0107
Education	0.0845***	0.0141	0.0256***	0.00753	0.0296***	0.00868
All	0.0431**	0.0192	0.0111	0.00739	0.0144	0.00901

*** p<0.01, ** p<0.05, * p<0.1

Controls (all variables relate to 2006): Dummies for education (college graduate, secondary school and vocational school), work experience in years, dummies for gender, municipality and firm size categories

Table 12: The effects of suspected cheating behavior on the changes of selected firm-level indicators in 2006-2007

(OLS regressions)

	Proxies of 'cheating' beavior					
Controls	P>0.5		w>MW		w>1.5 MW	
Change of average wage (log)						
No	0.134***	0.00860	0.0775***	0.00636	0.0875***	0.00668
Yes	0.0778***	0.00992	0.0365***	0.00740	0.0432***	0.00750
Change of employment (log)						
No	-0.0329***	0.00907	-0.0274***	0.00645	-0.0314***	0.00688
Yes	-0.0262***	0.00960	-0.0196***	0.00681	-0.0244***	0.00725
Change of sales revenues (log)						
No	-0.0640***	0.0152	-0.0242**	0.0113	-0.0284**	0.0128
Yes	-0.0499***	0.0167	-0.0151	0.0120	-0.0170	0.0136

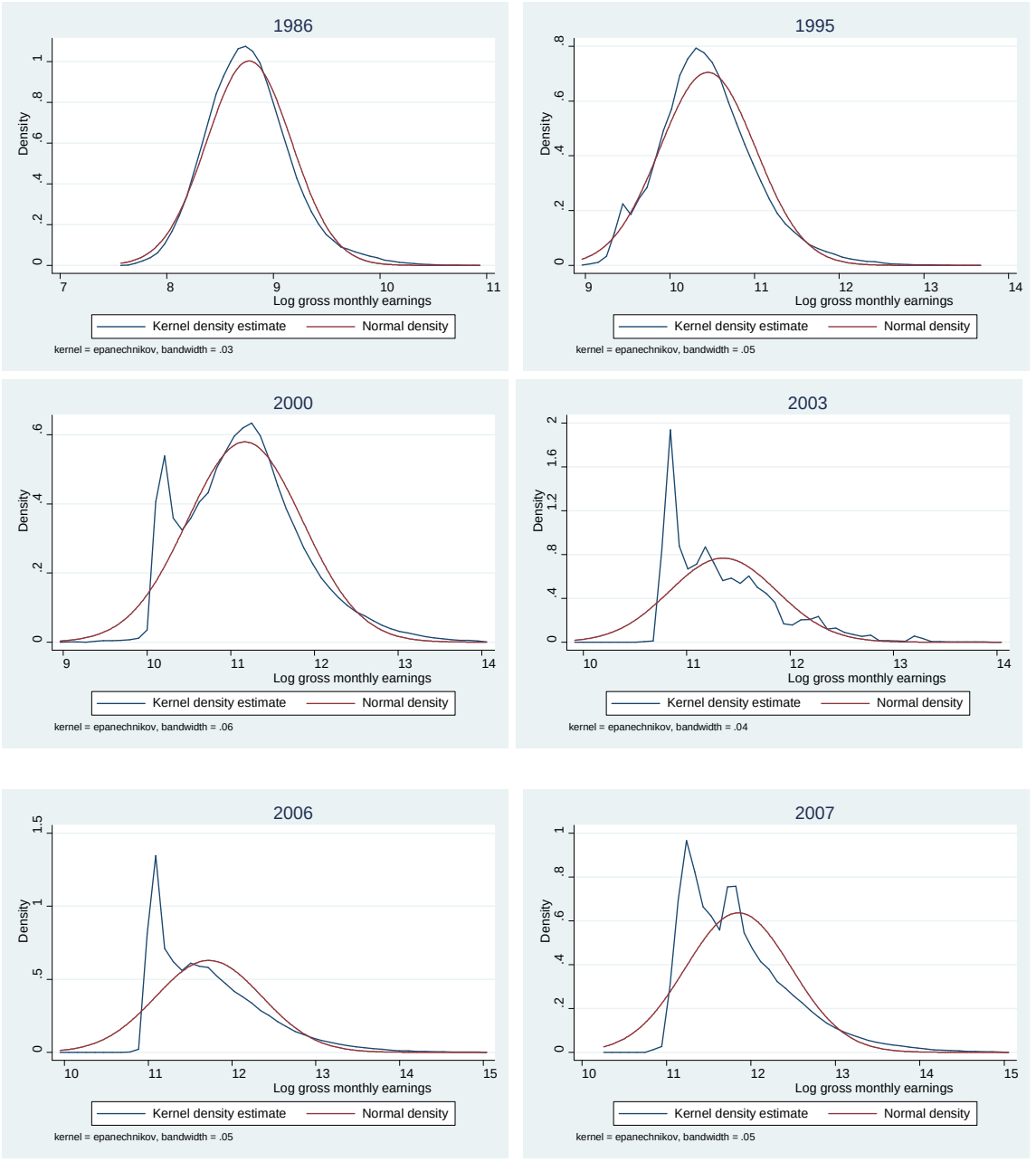
*** p<0.01, ** p<0.05, * p<0.1.

Sample: Firms observed in the Wage Survey in 2006 and 2007. Number of observations 4150 except for sales revenues (4173).

Controls include skill shares, average wage, average age and dummies for sectors, regions, type of municipality and state ownership

Figure 1: The wage distribution in Hungary, 1986-2007

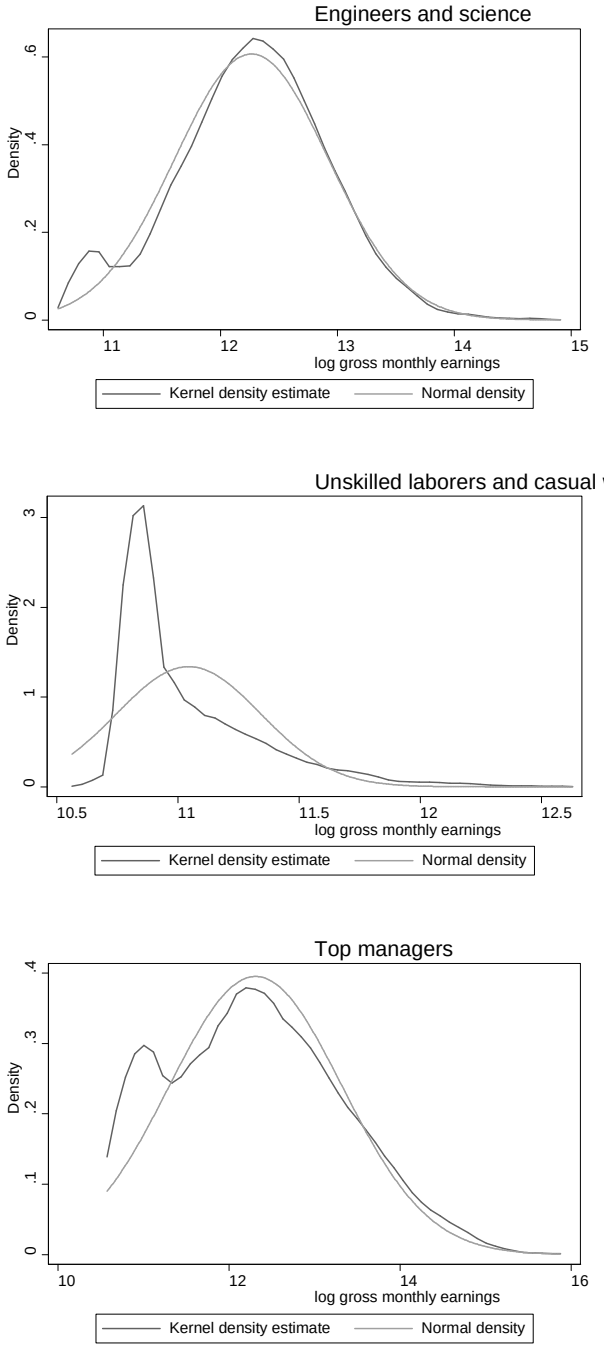
Kernel density estimates for gross monthly earnings in the private sector



Source: Wage Survey, private sector (firms)

Figure 2: The size distribution of earnings in three occupational groups in 2003

Private sector, full-timers, Wage Survey 2003



Note: The logarithm of the minimum wage was 10.82 in 2003

Figure 3: Wage distribution before and after the transformation
Enterprise sector, full-timers, Wage Survey 2003

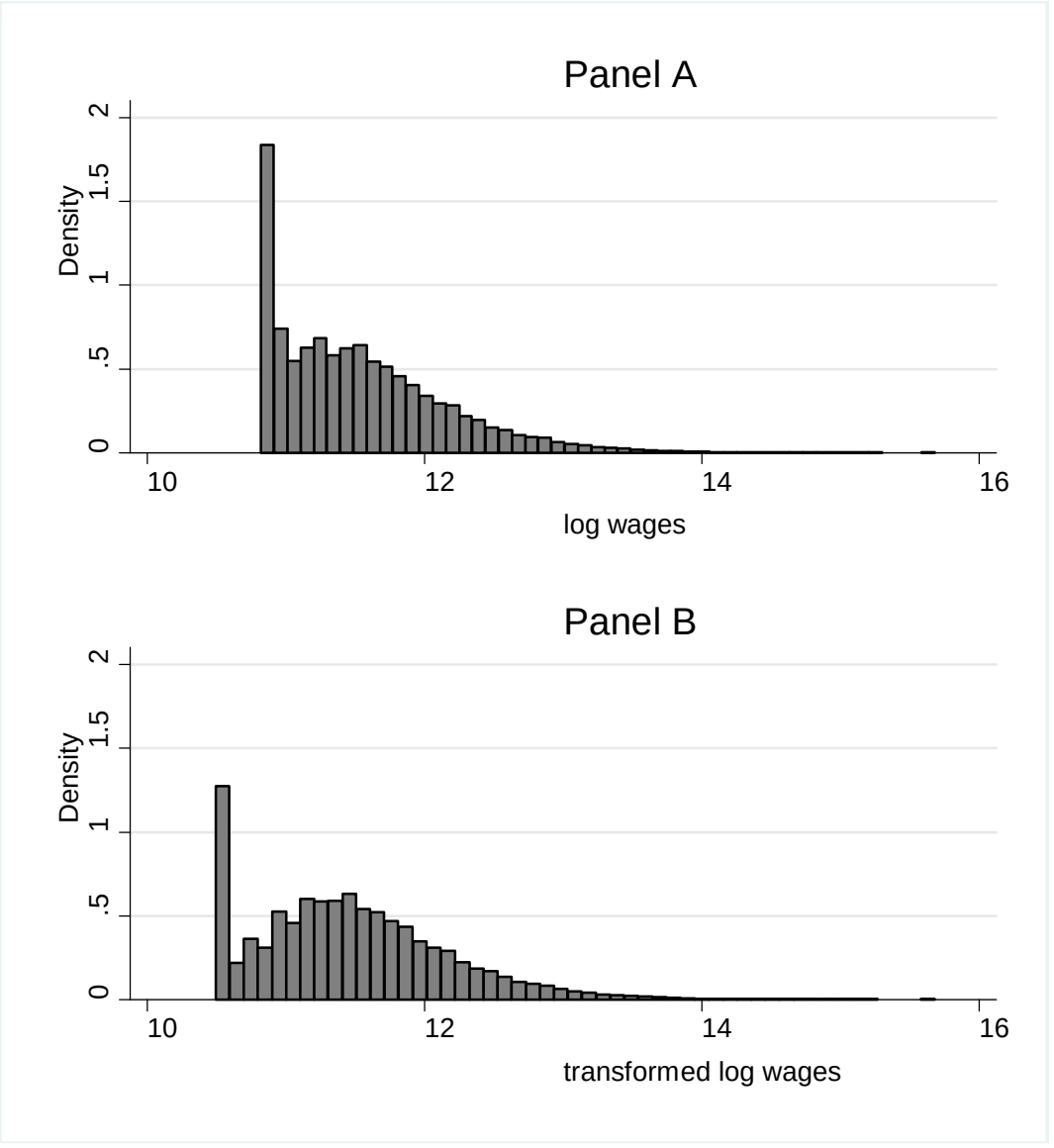
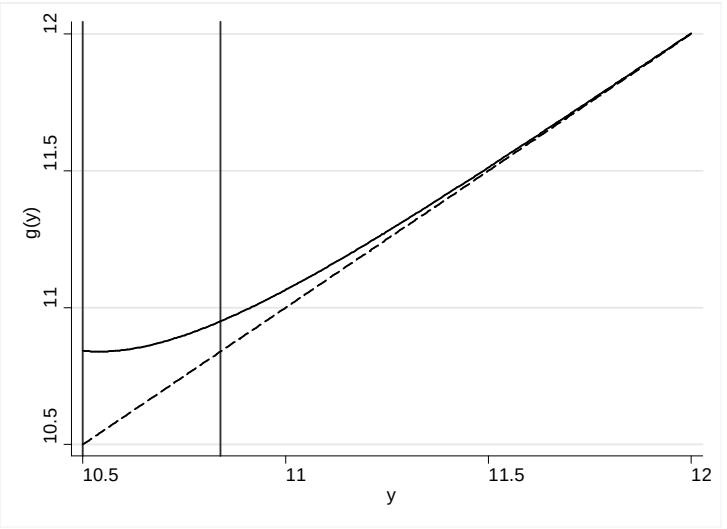


Figure 4: The shape of g for the parameter choice r = 0.3



Appendix 1

Table A1 Occupational classification used in the double-hurdle model

Occupations	Type*	Definition (based on standard classification of occupations)
Agricultural	E	Codes 61-64 and 92 comprising the drivers of agricultural vehicles
Construction	S	Code 76
Service	S	Codes 52-53 except 532, 533 and 536. Includes transport, mail and telecommunication
Trade	S	Codes 51 and 421, 422 and 429 comprising cashiers
Industrial	S	Codes 71-75
Other blue-collar		
Cleaners	E	Code 911
Unskilled laborers	E	Codes 913-919
Machine operators	E	Codes 81-83. Includes the operators of mobile machines such as cranes,
Porters and guards	E	Codes 912 and 536 comprising porters and security guards, respectively
Drivers	S	Code 833, 835, 836 Car, truck and bus. Excludes the drivers of agricultural vehicles
White-collar		
Office clerks	W	Codes 41-42 and 532-533 comprising office based jobs in health and social services
Technicians, assistants	W	Codes 31-34
Administrators	W	Codes 35-39
Managers	W	Codes 11-14
Professionals	W	Codes 21-29 except 22-24 (teachers and doctors)
Teachers and doctors	W	Codes 22-24
Armed forces	W	Codes 01-03

* E: elementary; S: secondary; W: white-collar

Table A2: Selection to the estimation sample of Test 1 – Probit marginal effects
Dependent variable: 1 if the firm is observed in both 2000 and 2003, and has non-missing data
Sample: Firms employing 5-20 workers observed in WS 2000

	Marginal effect	Z-value
Number of employees	0.0009	0.87
Average wage	0.0153	0.85
Fraction earning less than Ft 40,000	-0.0238	0.87
Sales revenues fall short of material costs (dummy)	-0.0433	1.42
Profit/worker	0.0027	1.22
Profit/worker missing	0.1354	1.86
Fixed assets per worker	-0.0000	0.01
Fixed asstes per worker missing	-0.0346	1.30
Share of men	-0.0208	0.86
Share of workers with primary education	0.0489	1.32
Share of workers with vocational education	0.0523	1.77
Share of workers with college/university education	0.0146	0.35
Average age	0.0007	0.51
Micro-region unemployment rate (log)	0.0116	1.16
Number of observations (small firms in the 2000 wave of the WS)	2,516	
Number of small firms in the estimation sample (panel 2000-2003)	263	
Observed P	0.104	
Predicted P	0.101	
Wald chi2 (significance)	26.65	0.2248
Pseudo R2	0.0146	
Joint significance of the sector dummies		0.2346

Note: Marginal effects at the mean of the independent variables. The coefficients of the one-digit sector dummies are omitted. All variables relate to 2000.

Table A3: Selected indicators of the estimation sample in Test 1

	All firms		Low-wage firms[a]	
	Cheaters	Non-cheaters	Cheaters	Non-cheaters
Number of firms	97	166	91	103
Average number of employees, 2000	13.8	13.0	13.8	13.0
Average number of employees, 2003	12.7	11.6	12.9	11.6
Average wage, Ft, 2000	46,146	87,385	44,582	53,509
Average wage, Ft, 2003	67,791	123,982	66,681	88,473
Fraction paid the MW, 2000	0.22	0.12	0.24	0.19
Fraction paid the MW, 2003	0.48	0.12	0.50	0.19
Fraction earning less than Ft 40,000, 2000	0.59	0.32	0.63	0.52
Fraction earning less than Ft 50,000, 2000	0.72	0.43	0.76	0.65
Fraction unskilled, , 2000[b]	0.16	0.13	0.16	0.18
Fraction unskilled, 2003[b]	0.15	0.11	0.15	0.14
Agriculture	0.07	0.05	0.08	0.08
Industry	0.24	0.25	0.22	0.21
Constructuction	0.14	0.13	0.15	0.17
Trade, hotels and restaurants	0.36	0.32	0.35	0.31
Other sectors	0.15	0.23	0.15	0.25
Mean log changes(st. dev.) 2000-2003				
Real average wage	0.26	0.25	0.27	0.35
	(0.26)	(0.34)	(0.26)	(0.36)
Residual wage	0.06	0.04	0.08	0.12
	(0.25)	(0.30)	(0.23)	(0.32)
Employment	-0.11	-0.17	-0.09	-0.19
	(0.39)	(0.46)	(0.38)	(0.51)
Share of unskilled workers (pct points)	-0.97	-2.31	-1.35	-3.86
	(0.14)	(0.17)	(0.14)	(0.20)

a) Firms with at least one worker earning less than Ft 40,000 in 2000.
b) Workers with complete or incomplete primary school attainment

Table A4: Firms suspected of wage under-reporting on the basis of DH estimates (2003, per cent)

Criteria applied:	All firms=100 (N=263)	Firms with at least one MW earner = 100 (N=194)
P>0 for at least one worker	35.4	65.9
w>MW for at least one worker	38.0	70.9
w>1.5 MW for at least one worker	31.1	58.1

Note: *w* stands for the simulated 'genuine' wage from the DH model of 2003. *P* stands for the estimated probability of wage under-reporting.

Figure A1: The distribution of individual wages in 2003 in small firms employing at least one MW earner

Sample: Firms employing MW earners in the estimation sample of test 1

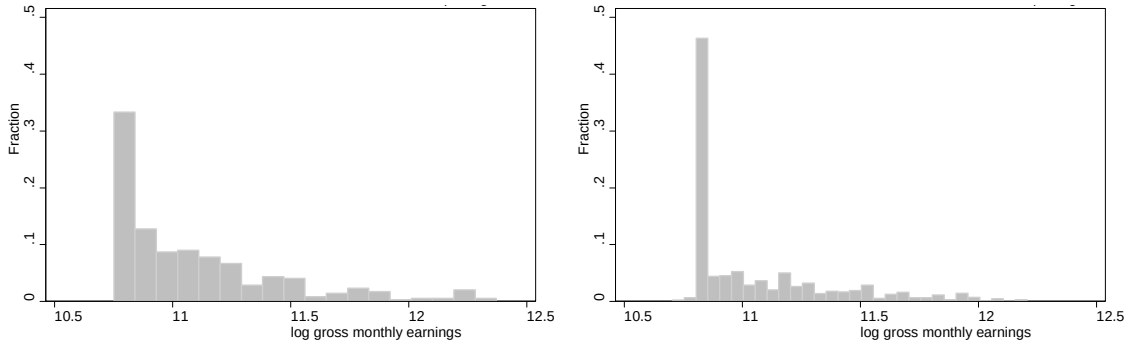


Table A5: Selection to the estimation sample of Test 2 – Probit marginal effects

Dependent variable: 1 if the employer is observed in both 2006 and 2007, 0 otherwise

	Marginal effect	Z-value
Number of employees	0.0141	18.9
Log gross wage	0.0182	6.24
Profit/worker	3.98e-07	0.00
Fixed assets per worker	3.98e-07	0.00
Men	0.0329	2.21
Primary education	0.03606	7.81
Vocational education	0. 05130	15.42
College/university education	-0. 01983	-4.77
Age	0.0292	21.84
Number of observations (employees in the 2006 wave of the Wage Survey)	166,949	
Number of employees in the estimation sample (panel 2006-2007)	56,515	
Observed P	0.338	
Predicted P	0. 371	
Wald chi2 (significance)	921.2	0.000
Pseudo R2	0.0458	
Joint significance of the sector dummies		0.000

Note: Marginal effects at the mean of the independent variables. The coefficients of the one-digit sector dummies are omitted. All variables relate to 2006.

Appendix 2

Minimum wage regulations in Hungary

Target and coverage. A single national monthly gross minimum wage was introduced by Hungary's last communist-led government in 1989. The minimum wage relates to monthly pre-tax base wages, that is, total monthly earnings net of overtime pay, shift pay and bonuses. Starting from 2007 weekly, daily and hourly levels are determined, too. The minimum wage is legally binding and covers all wages, including those paid to the self-employed by their own businesses. For part-timers, who account for about 5 per cent of total employment, the wage floor is proportionately lower. In 2006-2008 further minima applied to skilled workers (1.25MW) and young skilled workers (1.2MW). In 2009 the minimum for young skilled workers was abolished.

MW setting. The minimum wage is negotiated in a consultative body of employers and unions (Council of the Reconciliation of Interests). The government usually steps into the process at the end, by accepting the recommendations of the Council, but it is authorised to make a unilateral decision in case the negotiations fail, as it happened in 2001.

Level of the MW. At its introduction the MW amounted to 35 per cent of the average wage (AW), while in 2000 it stood at 29 per cent. Viktor Orbán's first government (1998–2002) nearly doubled the MW, by raising it from Ft 25,500 in December 2000 to Ft 40,000 in January 2001 and Ft 50,000 in January 2002. The two hikes raised the minimum wage–average wage ratio to 39 per cent and 43 per cent, respectively. Since 2003, the MW/AW ratio slightly fell but remained above its pre-hike level. See Figure A2.[18]

Compliance. The Wage Survey's data suggest that sub-minimum wages accounted for less than 1 per cent of all wages in each year since 1989. Estimates based on personal income tax reports and pension contributions hint at higher rates, but these data do not allow proper adjustment for hours worked during the year or daily.

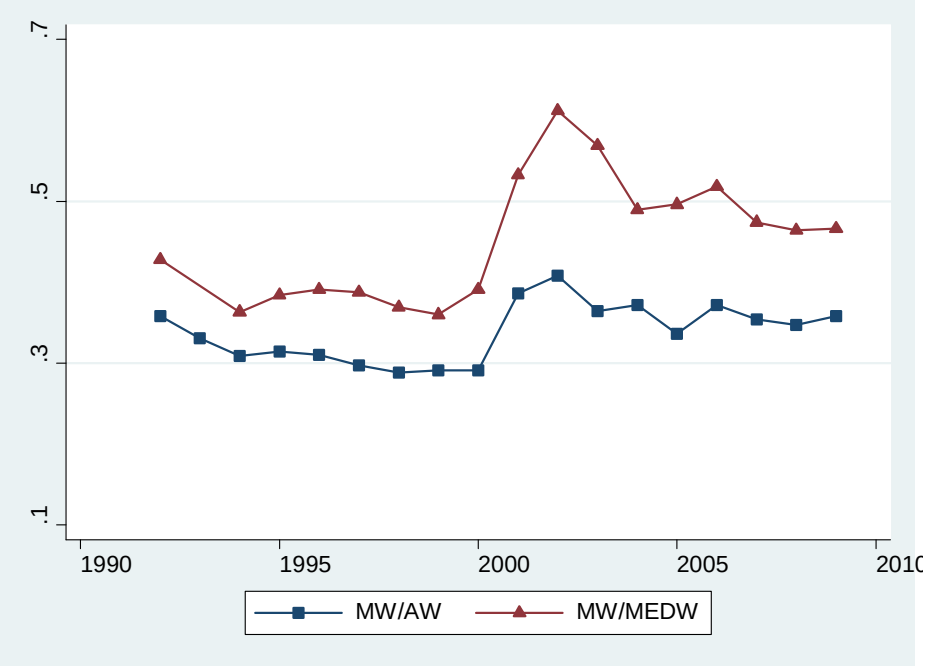
Fraction of employees affected. The fraction of workers paid 95–105 per cent of the MW amounted to 5 per cent in 2000. It jumped to 19 per cent in May 2002 in firms employing five or more workers and increased substantially in larger firms, too (Figure A2, bottom panel). The ratio fell to 10–12 per cent by 2004 and fell further substantially after 2006, when the tax authority started to interpret MW payment as a signal of wage under-reporting.

Taxing the MW. In 1989-2001 the MW was subject to linear social security contribution and progressive personal income tax. In 2002 it became free of personal income tax. In 2007, a minimum social security contribution base amounting to 2MW was introduced, as discussed in Section 3 of the text. This measure was abandoned in 2011.

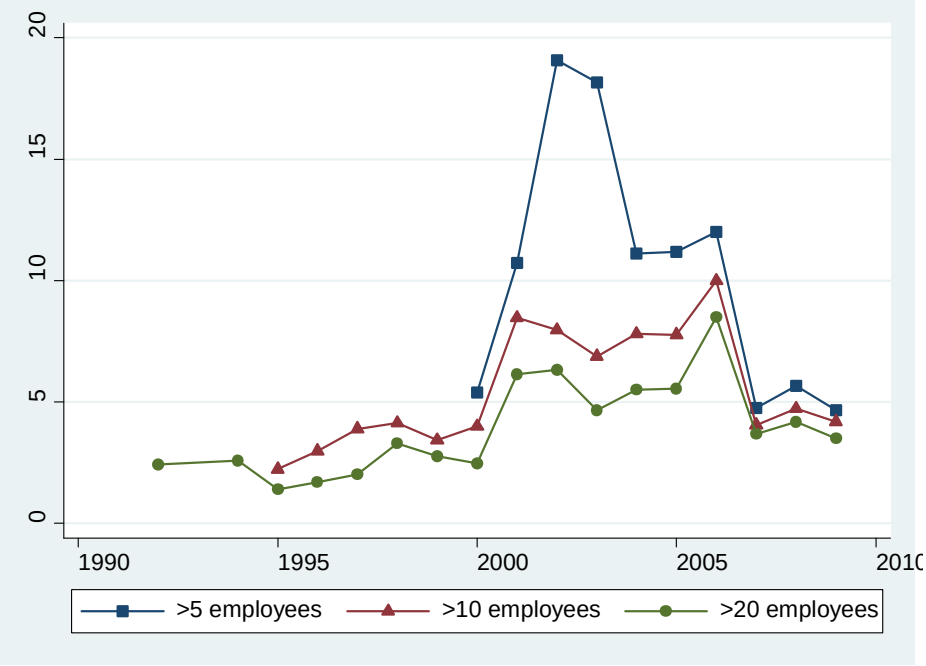
[18] All data quoted in this Appendix come from the Wage Survey.

Figure A2

(a) The MW compared to the average wage and the median wage 1992-2009



(b) Fraction paid 95-105 per cent of the MW, 1992-2009



Source: Wage Surveys. The survey was extended to smaller firms in 1995 and 2000