Comment on "Unemployment Compensation and Wages: Evidence from the German Hartz Reforms" by Stefan Arent and Wolfgang Nagl

Ludsteck, Johannes and Seth, Stefan

IAB Nürnberg (Institute for Employment Research Nuremberg)

8 November 2013

Comment on
"Unemployment Compensation and Wages: Evidence from the German Hartz Reforms" by Stefan Arent and Wolfgang Nagl

Johannes Ludsteck and Stefan Seth[*]

Institute for Employment Research (IAB) Nürnberg

November 8, 2013

Abstract

Arent and Nagl (2013) use the BA Employment panel 1998-2007 to identify effects of the German Hartz reform and find that it caused a considerable reduction of wages. Our replication study suggests that their clear and strong conclusions are based on implausible assumptions regarding the error structure of their regression models and on a too coarse modelling of the time effects. They become blurred and weak once better estimates of the standard errors are obtained and the development of wages is investigated at a finer time grid. Furthermore, Arent and Nagls' reform effects shrink considerably when a more appropriate price index is used to deflate the wages and when the censoring of wages is treated correctly. Methodological considerations suggest that their conclusions depend on several further daring and untested assumptions.

JEL classification: J08; J31; J65
Keywords: Hartz reforms; unemployment compensation; wages

1 Introduction

In an interesting and ambitious paper, Stefan Arent and Wolfgang Nagl (2013) analyze the effect of a decrease in unemployment benefits on wages. Their theoretical argument goes that lowering workers' outside options will depress wages. They employ the Hartz IV reform as a natural experiment and claim to find strong evidence that decreased unemployment compensation has an adverse effect on wages.

We take issue with their finding for the following reasons. A fundamental objection to Arent and Nagl's approach follows from the fact that they cannot use a standard control/treatment group approach because the Hartz reforms

affected *all* German employees.[1] Therefore, it would have been necessary to ensure not to confuse a supposed Hartz effect with the impact of other macro-economic factors. To our judgement, the authors failed to do so. By pointing to endogeneity problems, they omit central determinants of wages from their model, as e.g. the unemployment rate. Instead they include the industry level gross value added (per employee) in their analysis, which seems problematic too because it is *per definition* a function of individual wages, the dependent variable in their regressions.

A more detailed inspection of their approach reveals several further problems that put the reliability and precision of their results into question. First, we find that deflating the wages with a more appropriate price index decreases their Hartz effect estimates considerably. Second, the precision of their results seems to be overstated by an order of magnitude due to ignorance of clustering and serial correlation of the residuals. Third, Arent and Nagl do not properly address the problem that the wage variable in the data is right censored. This would be important especially when looking at the subpopulation of highly qualified employees, however. Fourth, modelling the development of wages at a finer time grid reveals a downward trend of wages starting in 2004 – well before the Hartz IV reform, which became law in 2005.[2] So the time line of events alone does not allow to make a causal claim here.

2   Arent and Nagl's model and slight extensions

Arent and Nagl (2013) employ the BA-Emplomynent Panel 1998–2007 (but use only the years 2000–2007 in their regression samples) to estimate wage regressions of the form

$$\ln(w_{it}) = \beta_0 + \beta_1 \, LUA_t + controls + a_i + u_{it} \tag{1}$$

where w_{it} denotes real monthly wages of indivual i in quarter t, LUA_t is a period dummy taking on value 1 for the years 2005–2007 (and zero otherwise), *controls* contains individual, establishment and industry level control variables,[3] a_i denotes an individual specific fixed effect, and u_{it} denotes a residual term.

Arent and Nagl emphasize that the Hartz reform generates quasi-experimental conditions and argue that purging the wages from the control variables in their model allows to interpret the coefficient of the *LUA* dummy ('Lower Unemployment Assistance') as a measure for the effect of lowering the unemployment assistance on wages.

We find it difficult to follow this argument due to the absence of a control group. Experimental and quasi-experimental designs usually compare the development of a group affected by a reform (the treatment group) with that of another group (the control group) *that is subject to the same common influences except the reform*. This allows the researcher to ignore influences that

---

[1]For example, the choice of an appropriate deflation measure would be a minor issue if a German control group were available.

[2]the integration of the unemployment insurance with the social assistance benefits took place in 2005 and the reduction of the maximum unemployment entitlement period became effective in February 2006.

[3]The controls include age, age squared, professional status, firm size, the firms's age structure, individual job tenure, annual values of the industry-specific gross value added per worker and dummy variables for each quarter.

affect *both groups in the same way*. Since a control group is absent here (*all* German employees are affected by the reform), controlling for macro level impacts is required.

Arent and Nagl are aware of the problem and try to tackle it by including the industry level gross value added per worker as a macro level control. The gross value added appears problematic as a regressor, however, since it is defined as a function of wages[4] and therefore must be highly correlated with the dependent variable *by construction*. Arent and Nagl emphasize that controlling for a number of individual and firm level characteristics helps to create quasi-experimental conditions. But their controls are almost time-invariant (the skill dummies[5]) or change steadily over time (the age profile, firm size dummies and the firm's age structure), and therefore cannot capture the effects of *discrete* changes of *macro* shocks (as e.g. export demand).

Finally, in order to obtain consistent estimates of the precision of the Hartz dummy coefficient, the multilevel structure of the model (it includes both micro and macro level regressors) has to be taken into account. A closer look at the error structure of the model suggests that the coefficient standard deviations reported by Arent and Nagl might be underestimated by an order of magnitude since aggregate shocks that are not represented in the regression function may generate clustering of the residuals at the year level. Expressed formally, the residual $u_{i,t}$ in model (1) should be written as $u_{it} \equiv v_{i,t} + \eta_t$ with fixed time effect η_t. Betrand et al. (2004) demonstrate that serial correlation of the residual component $v_{i,t}$ creates over-rejection problems that appear notoriously in Difference-in-Differences estimates and recommend a moving blocks bootstrap to tackle it. Though model (1) is not a Difference-in-Differences design, the problems are similar or even worse here.

Even if one provisionally accepts Arent and Nagl's assumptions regarding the exogeneity of the regressors and the error structure, a causal interpretation of the Hartz dummy coefficient becomes questionable if the time pattern of purged wages is inspected directly. A finer time grid is obtained by replacing the reform dummy *LUA* by a set of time dummies for the years 2003 to 2007.[6] This yields the following model

$$\ln(w_{it}) = \beta_0 + \sum_{\tau=2003}^{2007} \gamma_\tau \, D_t^{(\tau)} + controls + a_i + u_{it} \tag{2}$$

where $D_t^{(\tau)}$ is a year dummy for year τ. It takes on value 1 for all quarters of year τ and zero otherwise.

---

[4]The GVA is computed as output at market prices minus intermediate consumption. Clearly, wages constitute a large share of the value added.

[5]See our remark in the appendix.

[6]The year dummies for 2000 to 2002 are omitted to avoid collinearity problems with the age terms. Extending the base improves the estimation of the age/trend component. Deviating from Arent and Nagl (they include the age variable at yearly precision in their models), we use the age at quarterly precision.

3 Results

3.1 Standard error issues

As mentioned above, the standard errors reported by Arent and Nagl may be underestimated by and order of magnitude due to the neglect of clustering and serial correlation of the residuals. Strong serial correlation of the residuals can be shown by regressing the estimated residuals $\hat{u}_{i,t}$ on their lags $\hat{u}_{i,t-1}$. This yields a quite high correlation of 0.8. (All results presented in our comment refer to western German men.) That the estimated Hartz effect (the point estimate, which can be found in row 1, column 1 of Table 1, is -0.024) seems to be much less precise than suggested by the tiny standard error (0.0002) reported in Arent and Nagl's paper, can also be demonstrated by running separate regressions based on the observations from one quarter only. They yield the reform dummy estimates -0.018, -0.022, -0.027 and -0.029 for the first, second, third and fourth quarter, respectively. Considering that the dummy coefficient measures an average over the three years 2005, 2006 and 2007, it is quite irritating that the 'Autumn' Hartz effect exceeds the 'Winter' effect by roughly 60 percent.

More sensible estimates of the standard errors can be obtained from a moving blocks bootstrap procedure (as recommended by Bertrand et al.). If blocks containing all observations from 2, 4 or 6 successive quarters are drawn randomly with replacement from the data, we obtain standard errors of 0.008, 0.011 and 0.014, respectively (see columns 3–5 in row 1 of Table 1). This exceeds the figure reported by Arent and Nagl by a factor of roughly 40 to 70. Even these much larger standard errors may overstate the precision of the estimates, however, since the short time dimension of the data (32 waves only) puts tight limits on the length of the bootstrap blocks.[7]

3.2 The role of the price index

A further issue in Arent and Nagl's study regards the choice of the price index. They use the harmonized price index for Germany to deflate the wages. Since the harmonized index is developed mainly in order to improve the comparability of price changes *across* European countries, it is not completely representative for Germany. Due to the absence of a control group, the choice of the price index may influence the results considerably. A simple comparison of the harmonized price index with the national German consumer price index shows that this is the case. Whereas the development of both indexes was almost identical in the years 2000 to 2004, the growth of the harmonized index exceededs that of the national index by 0.51 percentage points. A simple rough computation suggests that Arent and Nagl would have obtained a considerably smaller Hartz effect of $-2.4 + 0.51 \approx -1.9$ percent if they had employed the more appropriate index. This is confirmed by the considerably smaller Hartz dummy coefficient estimates in rows 3 and 4 of Table 1 that are obtained from our own preparation of the BA panel. A closer inspection reveals that the price index is the main source of

---

[7]Bertrand et al. consider a design with several treatment groups (U.S. States). Their blocks comprise the *entire time series of the observations of a state.* Since we have only one treatment group, this is not feasible for our application. Bertrand et al. note that their bootstrap procedure performs well only if the number of states is large enough (at least 20). See Fitzenberger (1997) for an expositon of the moving blocks bootstrap.

Table 1: Hartz dummy coefficient estimates, obtained from estimating models (1) for western German men. Dependent variable: log real wages.

| Model/Sample | Point Estimate | Asy.[a] | Standard Error Estimates Moving Blocks Bootstrap | | |
| | | | 2 Waves | 4 Waves | 6 Waves |
	(1)	(2)	(3)	(4)	(5)
A. Wages Deflated with the Harmonized Consumer Price Index					
A & N's Sample	-0.0240	0.0002	0.0077	0.0114	0.0143
A & N's Sample, model includes only age, age^2 and quarter dummies	-0.0232	0.0002	0.0074	0.0109	0.0117
B. Wages Deflated with the National Consumer Price Index					
B.I All Skill Groups Together					
Include All Censored Obs.	-0.0166	0.0003	0.0053	0.0076	0.0088
Drop All Censored Obs.	-0.0142	0.0003	0.0041	0.0053	0.0064
Fixed Effects Tobit	-0.0077	–[b]	–	–	–
B.II Separate Regressions by Skill Group (Fixed Effects Tobit Estimates)					
Low Skilled	-0.0093	–	–	–	–
Medium Skilled	-0.0073	–	–	–	–
Highly Skilled	-0.0054	–	–	–	–

Data Sources: BA-Employment Panel 2000–2007, data preprocessing by Arent and Nagl (rows 1 and 2), and by Ludsteck and Seth (rows 3–8).

The models in rows 3–8 omit the gross value added per employee. All bootstrap standard errors are obtained from 100 bootstrap replications.

Notes:

[a] 'Asy'. refers to asymptotic standard errors that account for heteroscedasticity.

[b] Bootstrap standard errors are not computed due to prohibitively high computational costs.

the deviations between their and our results.[8]

3.3 The censoring problem

Further inspection of the data suggests that the Hartz dummy coefficient may be biased by improper treatment of the wage censoring. As is well known, the wage of the employment register data (Beschäftigtenstatistik) are right-censored at the earnings limit for social security contributions ('Beitragsbemessungsgrenze'). The censoring share is moderate for the entire sample, about 5 to 10 percent, slightly varying by year. For the highly qualified men in western Germany, it amounts to roughly 50 percent, however.[9] This may bias standard OLS and fixed effects regression models severely and therefore requires proper treatment, e.g. by using Tobit models.

Arent and Nagl try to tackle the censoring problem heuristically by truncating the sample at the 5th and 95th percentile of the unconditional wage distribution. This is problematic for several reasons. First, the application of standard OLS or fixed effects estimators to a truncated sample delivers still biased results. Since roughly 50 percent of the wages are censored for the highly qualified men in western Germany, comparing the results from standard OLS

---

[8]We obtain a Hartz effect of -0.0239 if the wages are deflated with the harmonized price index in our data.

[9]See e.g. Büttner and Rässler (2008) for statistics on the censoring shares.

or fixed effects regressions for this group with the unskilled workers (where censoring is below 2 percent), appears to be problematic, to say the least. Second, Arent and Nagl's heuristic truncation procedure leaves the majority of the censored observations in the sample if the share of right-censored observations exceeds 5 percent, which clearly is the case for some subsamples (e.g. western German men and the highly qualified employees). Due to an error in their pre-processing scripts, Arent and Nagl actually drop only about 1 percent of the top observations instead of 5 percent as reported in the paper. Third, the censoring limit was increased by roughly 10 percent in 2003.[10] Such year-to-year changes may translate directly into artificial changes of period or year or dummy coefficients if they are not tackled appropriately using Tobit models. Fourth, truncating five percent of the wages at the bottom might drop a good deal of low but valid wages. It deviates from the standard practice to disregard wages of full-timers that are smaller than 100 or 200 percent of the marginal part-time income threshold.[11]

In order to inspect the sensitivity of Arent and Nagl's results to censoring problems, we compute some further estimates of the Hartz effect. The first two are based on standard fixed effects regression models. The first includes all censored observations, the second drops all censored observations. The third accounts for censoring using Honoré's (1993) fixed effects Tobit estimator (details on the implementation of Honoré's estimator are described in the Appendix). All three regressions are based on data from our own preprocessing due to lack of correct censoring indicators in Arent and Nagl's data.[12] Since the harmonized price index is clearly not the best choice, it is substituted by the standard consumer price index for our alternative estimates. The results of the Tobit regression are shown in row 5 of Table 1. Now the Hartz effect has decreased again in magnitude, to less than one percent.

The most surprising results in Arent and Nagl's study are derived from separate regressions by skill groups. Since these regressions are conducted even separately for 5 industries, their results cannot be summarized in one number per skill group. The responses of the highly skilled to the reform are, however, greater in all industries – according to their results. For example, the Hartz effects are -1.58 and -2.76 percent, respectively, for the low and highly skilled men employed in the manufacturing industry. When we apply the fixed effects Tobit estimator to the skill subsamples, our results point in the opposite direction. In line with expectations from theoretical reasoning, the effect (-0.51 percent) is now even slightly smaller for the highly skilled. If the relationship of the point estimates to their standard errors is similar to rows 1–4, all these effects are probably insignificant, however. We note that estimating Tobit regressions for samples with roughly 50 percent censored observations appears daring. Therefore we would hesitate to believe in their results seriously even if all other problems mentioned above were absent. They are, however, good enough to show that Arent and Nagl's results on the highly skilled should not

---

[10]The increase was induced by the Beitragssicherungsgesetz (BSSichG), see Bungesgeset-zblatt, Jahrgang 2002 Teil I, Nr. 87.

[11]Less than 3 percent of the full-time spells are dropped even if the 200 percent threshold is used. Note furthermore that the marginal part-time threshold was raised considerably from 325 to 400 Euro in 2003.

[12]The treatment of *low* wages in our data preprocessing follows the standard procedure to drop wages not exceeding 200 percent of the marginal part-time income threshold.
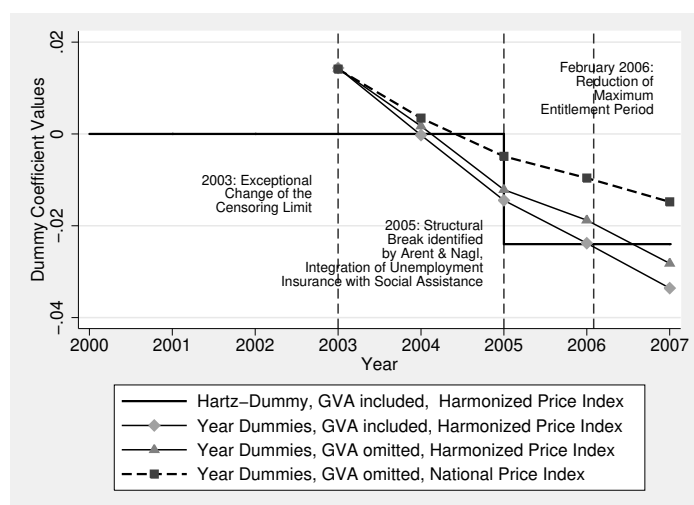
Figure 1: Hartz dummy coefficient estimates, obtained from estimating models (1) and (2) for western German men. Dependent variable: log real wages.

Data Sources: BA-Employment Panel
data preprocessing by Arent and Nagl for all results based on the harmonized price index, data preprocessing by Ludsteck and Seth for the results based on the national price index.
Legend: 'Hartz-Dummy' represents Arent and Nagl's original Hartz-Dummy effect from model (1), 'Year Dummies' represent estimates of the year dummy coefficients for model (2). 'GVA included' and 'GVA omitted' indicate whether the gross value added per worker is included or not as a control variable. All estimates are obtained using standard fixed effects models.

be taken seriously.

3.4 Reform effect estimates at a finer time grid

Since a cause should precede its effects according to common notions of causality, it appears important to check whether the Hartz reform effect really looks like the step function (it is represented in Figure 1) generated by Arent and Nagl's reform dummy specification.

We do this by estimating model (2) for wages deflated with the harmonized price index as well as with the national index. Though the results in Table 1 suggest that the standard least squares fixed effects estimates are slightly biased, they are employed here for the sake of simplicity. The results from this exercise are represented by the graphs in Figure 1. We find a wage decline that starts already in 2004 and gives no indication of a clear-cut downward jump afterwards. If the wages are deflated with the national price index, the graph becomes even flatter in the years 2006 and 2007. One can still argue that the wages decreased *cet. par.* in the years 2005–2007, but it is hard to imagine that a researcher who didn't know when the reform took place, would find a pronounced reform effect in the years 2005 to 2007 based on the year dummy coefficients.[13]

---

[13]Arent and Nagl find a structural break in 2005 based on a Chow test (see page 454 of their

In principle, Appendix Table A4 in Arent and Nagl's paper contains already enough information to put their interpretation of the Hartz dummy into question. They run placebo regressions to check the validity of their hypothesis. The placebo specifications are obtained by shifting the start of the Hartz period dummy successively to 2004, 2003 and 2002. Whereas the coefficient of the Hartz dummy starting in 2005 is negative (-0.024), shifting the start backwards to 2004 and 2003 and 2002 renders the coefficients positive (see Table A4 in their paper). Their values are -0.008, 0.0266 and 0.0213 for the respective years 2004, 2003 and 2002 (all significant at the 1 percent level according to the standard errors reported by Arent and Nagl). Combining these estimates allows to retrieve the pattern of year-specific deviations from the overall trend and clearly shows that the downward movement started already before 2005. In general, highly significant placebo effects should *not* be taken as evidence of the validity of an experimental design, even if they have the 'wrong sign.' Effective placebos rather indicate the presence of ignored confounding factors.

Finally, to be perfectly clear in this important respect: The fact that we present results from equations (1) and (2) does not mean that we consider them as valid or appropriate models to evaluate the reform effects. Considering the bootstrap standard errors in Table 1 the significance of the year dummy coefficients is questionable, and exploratory regressions revealed that a longer estimation period would be required to improve the joint estimation of the age/trend profile and the year dummies.[14] The extended model regressions are conducted only in order to show that Arent and Nagl's coarse step function gives a keyhole view on the issue to be investigated and may attribute wage changes to the reform that actually took place in advance.

4 Conclusion

Arent and Nagl have, to their merit, raised an important empirical issue: What was the effect of the Hartz reforms on wage formation? They try to answer it based on simple fixed effects regression models of wages on several controls and a post-reform period dummy. Due to the absence of a control group, however, their ceteris paribus interpretation of the reform effects is rather assumed than ensured by appropriate control for relevant confounding factors at the *macro* level. Believing in their ceteris paribus clauses is a matter of faith and seems to be daring in several respects. Moreover, the Hartz effects depend on which price index is used to deflate the wages, and shrinks considerably (from about 2.4 percent to 1.6 or 1.4 percent for western German male employees) if the harmonized prize index is replaced by the more appropriate national index. If the

---

article). Unfortunately they do not report any further details of the test procedure. When we performed standard Chow tests (by interacting all regressors with the period dummy and testing for joint significance of all interaction effects) and based the test on the standard errors returned by Stata, we obtained structural breaks for every year. Arent and Nagl try to corroborate their finding by running additional pooled regressions based on sub-periods before (2003) and after the reform (2007). It is, however, completely unclear what to learn from these regressions. They are based on *linear* models (see Appendix Tables A1 and A2) whereas the regressions in the main part are *log-linear*. If the true model is log-linear, then the linear one must be misspecified and structural breaks may appear artificially or be hidden as a consequence of neglected nonlinearity.

[14]The joint identification of age/trend and year effects is further hampered by Arent and Nagl's choice to use a highly balanced panel.

computation of the standard errors is based on more general and plausible error structures, the significance of the effects becomes unclear. Slight extensions of Arent and Nagl's period dummy specification suggest that wages started to decrease before the integration of unemployment security with social assistance benefits in 2005 and the reduction of the maximum entitlement period that became effective in February 2006. An assessment based on notions of causality that are commonly applied in empirical research makes it difficult to attribute the observed effect uniquely to the Hartz reform. Expressed in simple terms: Arent and Nagl's keyhole view suggests strong contours which melt away once we open the door to get a panorama view. Given the problems with Arent and Nagl's analysis, their conclusion remains unconvincing.

When we account for the censoring of wages using fixed effects Tobit models, the reform effect shrinks again (to about 0.8 percent for western German men), and Arent and Nagl's finding that the reform reduced the wages of the highly skilled employees more than those of the medium and low skilled, is reversed. Though we fight shy of interpreting our results as reform effects, they appear more plausible in terms of models that allow for on-the-job search. With on-the-job search, wage offers are compared to the *prevailing* wage rate, rather than unemployment benefits. Given the relatively low unemployment rates of the highly skilled workers, on-the job search seems to be of more relevance for them than to the lower-skilled workers. This would account for the increasing wage gap between the better skilled workers and the less skilled workers we observe.[15] In contrast, the view suggested by Arent and Nagl would suggest, counterfactually, a narrowing wage gap between the highly skilled and the less skilled workers in the wake of the Hartz reforms.

References

Arent, S., and Nagl, W. (2013). Unemployment Compensation and Wages: Evidence from the German Hartz Reforms. *Jahrbücher für Nationalökonomie und Statistik*, 233(4): 450–466.

Betrand, M., Duflo, E., and Mullainthan, S. (2004). How Much Should We Trust in Differences-in-Differences Estimates? *Quarterly Journal of Economics*, 119(1): 793–796.

Büttner, T., and Rässler, S. (2008). Multiple Imputation of Right-Censored Wages in the German IAB Employment Sample Considering Heteroscedasticity. Discussion Paper 44/2008, IAB, Nürnberg.

Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the rise of West German Wage Inequality. *Quarterly Journal of Economics*, 128(3): 967–1015.

Fitzenberger, B. (1997). The Moving Blocks Bootstrap and Robust Inference for Least Squares and Quantile Regression. *Journal of Econometrics*, 82(2): 235–287.

---

[15]See e.g. Card et al. (2013)

Honoré, B. (1993). Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects. *Econometrica*, 60(3): 553–565.

Authors' address

Corresponding author: Johannes Ludsteck

Regensburger Str. 102, 90478 Nürnberg

email: johannes.ludsteck@iab.de, tel.: ++49/911/179-4810

A   Details on the implemenation of Honoré's fixed effects Tobit estimator

This section provides a brief description of the fixed effects Tobit estimation procedure.

Honoré's estimator is formulated for time-constant left-hand side censoring at zero only. It can be applied, however, without further changes to our data that show time-varying right-hand censoring, by applying the well-known transformation $\tilde{y}_{it} := c_t - y_{it}$ to our dependent variable $y_{it} \equiv \ln(w_{it})$ and rewriting the regression model in the form

$$\tilde{y}_{it} = c_t - x_{it}\, b - a_i - u_{it} \tag{3}$$

with artificial regressor c_t.[16] Then we obtain a regression model with time-constant left-censored dependent variable $\tilde{y}_{it}$.

Second, Honoré's estimator is based on time-differencing and therefore deviates slightly from the within-transormed fixed effects estimates that employ differences from individual-specific means. We use lag-4 differences $\Delta^{[4]} x_{i,t} \equiv x_{i,t} - x_{i,t-4}$. This induces a small loss information since 3 quarters are lost compared to the lag-1 differencing $\Delta^{[1]} x_{i,t} \equiv x_{i,t} - x_{i,t-1}$ which drops only one quarter. It is attractive, however, as it simplifies the estimation by eliminating seasonal effects and therefore allows us to drop the seasonal dummies from the regression model.

A third straightforward adjustment is required since the orthogonality conditions used to estimate the coefficients of Honoré's estimator are formulated for a single pair of waves only. To apply the estimator to a sample spanning more that two waves, we minimize the sum of the objective functions for several wave-pairs. To represent this formally, define $\tilde{Y}_t := (\tilde{y}_{1t}, \ldots, \tilde{y}_{Nt})'$ and $X_t := (x_{1t}, \ldots, x_{Nt})'$. If $\chi^2(\tilde{Y}_{t-4}, \tilde{Y}_t, \Delta^{[4]} X_t\, b)$ denotes the objective function for the wave pair $(t-4, t)$, the objective for the entire sample has the form

$$\Omega = \sum_{t=4}^{T} \chi^2(\tilde{Y}_{t-4}, \tilde{Y}_t, \Delta^{[4]} X_t\, b). \tag{4}$$

It is minimized using a derivative-free global optimization algorithm.[17]

Fourth, we omit the skill dummies 'unskilled blue collar worker,' 'skilled blue collar worker' and 'foremen' since they show extremely low time-variation. They change only in 0.22 percent, 0.22 percent and 0.04 percent (sic!) of all observations in Arent and Nagl's data and may therefore create optimization

---

[16]The 'coefficient' of c_t is restricted to unity in our implementation of the estimator.

[17]We use Nelder and Mead's flexible polyhedron search as implemented in the computer algebra system *Mathematica*. The algorithm was restarted 5 times with random starting values in the range $[-0.05; 0.05]$ in order to increase the likelihood to find the global minimum. The program code is available from the authors upon request.

problems in the Tobit models.[18] Of course, we checked whether these dummies matter as controls for the Hartz effect. Omitting them from Arent and Nagl's base regression (row 1 of Table 1) changes the Hartz dummy coefficient from -0.0240 to -0.0242.

---

[18]It is clear that the skill dummies are almost time-constant since the formal qualification is either acquired before persons enter the labour market, or it is acquired in an apprenticeship. In the second case, the change of the qualification is not visible in the estimation sample since then the change occurs at the end of the apprenticeship and apprenticeship spells are removed from the estimation sample.