ISSN 1503-299X

WORKING PAPER SERIES

No. 1/2018

Identification of the private-public wage gap*

Jørn Rattsø
Department of Economics
Norwegian University of Science and Technology

Hildegunn Ekroll Stokke
Department of Economics
Norwegian University of Science and Technology

Department of Economics

Norwegian University of Science and Technology
N-7491 Trondheim, Norway
http://www.ntnu.edu/econ/working-papers

* We appreciate discussions at the 12th winter school on inequality and social welfare theory in Canazei, the 2017 European Meeting of the Urban Economics Association in Copenhagen, the 2017 LAGV conference in Aix, the 2018 DEMOS-workshop on urban/labor economics, and the 2018 EALE conference in Lyon, and in particular comments from David Albouy, Nate Baum-Snow, Paolo Naticchioni, Kjell G. Salvanes, and Olmo Silva. We are grateful for the cooperation of Statistics Norway and funding from the Research Council of Norway (grant number 255509).
** Corresponding author.

Identification of the private-public wage gap*

Jørn Rattsø** and Hildegunn E. Stokke

Department of Economics, Norwegian University of Science and Technology

jorn.rattso@ntnu.no, hildegunn.e.stokke@ntnu.no

Abstract

Selection issues represent the main methodological challenge to the analysis of the private-public wage gap. Recent studies apply individual panel data and account for observable and unobservable characteristics including time-varying controls and worker fixed effects. Our starting point is a difference-in-difference model assuming that shift from public to private sector is a treatment. We reject the assumption of parallel wage trends between shifters and public stayers and suggest a new identification strategy where shifters early in the period studied are compared with workers still in the public sector that shift later. The estimates are based on rich register data for high-educated male workers in Norway during 1993-2010. Comparing three years before and after the shift year, the analysis shows that the wage gain from shifting to the private sector is about 10%. The identification strategy indicates that a difference-in-difference model comparing shifters with stayers leads to an overestimation of the wage gap by 20% (corresponding to 2 percentage points in our case). The counterfactual established with a comparison of early with late shifters gives a lower estimate of the additional private sector return.

Date: October 31, 2018

JEL codes: J31, J45, J62

Keywords: Private-public wage gap; Selection; Register data; Difference-in-difference

* We appreciate discussions at the 12th winter school on inequality and social welfare theory in Canazei, the 2017 European Meeting of the Urban Economics Association in Copenhagen, the 2017 LAGV conference in Aix, the 2018 DEMOS-workshop on urban/labor economics, and the 2018 EALE conference in Lyon, and in particular comments from David Albouy, Nate Baum-Snow, Paolo Naticchioni, Kjell G. Salvanes, and Olmo Silva. We are grateful for the cooperation of Statistics Norway and funding from the Research Council of Norway (grant number 255509).

** Corresponding author.

1. Introduction

The private-public wage gap is important for the understanding of public sector finances and the working of labor markets. Estimation of the size of the gap faces serious methodological challenges related to heterogeneity and selection. The understanding here is that public employees shifting to the private sector represent a treatment group. To estimate the treatment effect, the wage achieved in the private sector must be compared with the hypothetical wage obtained in the public sector, the counterfactual. The standard approach is to estimate the counterfactual based on the non-treated, that is, those staying in the public sector. We suggest a new identification strategy where shifters early in the period studied are compared with workers still in the public sector that shift later.

Recent overviews of public sector wage gaps, such as Giordano et al. (2014) and Lausev (2014), do not deal with the selection issues properly. The availability of register data over time allows for individual level panel analysis with correction of selection based on observables and unobservables. The identification of the gap can be based on shifters between the private and the public sector in models with worker fixed effects. Interesting applications include Bargain and Melly (2008) applying panel data in quantile regressions studying the wage distribution, recently extended by Bargain, Etienne and Melly (2016). Hospido and Moral-Benito (2016) estimate similar wage distributions for Spain. Rattsø and Stokke (2018) introduce the dynamic wage gap including different returns to experience in the two sectors. Schanzenbach (2015) adds variables representing ability of individuals as an alternative to worker fixed effects, notably an IQ-score. The worker fixed effect panel analyses handle selection based on time-constant unobservables, but selection issues remain. Alternative approaches include bound analysis introduced by Depalo (2017) and the estimation of a structural model as offered by Postel-Vinay and Turon (2007) and Dickson et al. (2014).

We use administrative data for Norway to estimate the private-public wage gap. The analysis concentrates on the core part of the public sector, public administration, with fairly high mobility towards the private sector. To limit the heterogeneity of workers further, we only include male high-educated workers (education beyond high school). The data cover the wage performance during 1993–2010 and include approximately 900,000 observations and 25,000

workers (of which 5,000 are shifters from the public to the private sector). Our starting point is the individual level panel model taking into account selection on observables and unobservables. We formulate a basic difference-in-difference model where shift to the private sector is the treatment and stayers in the public sector are the control group. Given the development of public sector wages, we study how a shift to the private sector affects the wage path. The specification includes each worker's wage three years before and after the shift year. The analysis rejects the assumption of parallel wage trends between shifters and stayers and shows that this comparison is invalid as a causal effect.

We suggest an identification strategy based on the assumption that shifters to the private sector are similar. The counterfactual is established comparing shifters to the private sector early in the period (early shifters) with workers who are currently in the public sector, but who shift to the private sector later in the period (late shifters). We are able to do this because we observe sector shifts over an extended period. The suggested strategy is related to the analysis of Borjas (2003) emphasizing workers entering the private sector relative to workers leaving the private sector. However, his interest is the effect of changes over time in wage compression in the public sector for wage differentials between 'entrants' and 'quitters'. We use the differences between early and late shifters to identify the effect of shifting from the public to the private sector.

Given identification based on comparison of early versus late shifters, the estimated private-public wage gap equals 10% (comparing three years before and three years after the shift year). Separating out the effect for each year, we find that the private wage premium is increasing over time, reflecting higher return to experience in the private sector. The basic difference-in-difference model comparing shifters with stayers overestimates the private-public wage gap by 20% (corresponding to 2 percentage points in our case). The overestimation bias represents a positive selection of shifters to the private sector compared to stayers in the public sector. It should be noticed that endogenous mobility and reverse causality remains another methodological challenge not pursued here.

As a check of robustness, and to exploit more data than in our analysis of shifters, we try out an alternative identification strategy exploiting that the recruitment to the public sector

differs across the business cycle. This can be understood as a difference-in-difference-in-difference model separating between private-public shifters recruited in booms versus recessions. The idea is that shifters that were recruited to the public sector during booms represent those with preferences to work in the public sector, and that are consequently most similar to the public stayers. The difference in return between those recruited in booms versus recessions gives a measure of the selection bias. The findings confirm the overestimation in the basic model comparing all shifters with the stayers.

The dataset and econometric model are described in section 2. Section 3 presents the basic difference-in-difference model comparing shifters and stayers. The identification strategy based on early/late shifters is applied in section 4. Section 5 discusses an alternative identification strategy based on the business cycle at time of labor market entry. Concluding remarks are given in section 6.

2. Data and econometric model

To handle the problems following heterogeneity of workers, we narrow down the comparison to male workers with higher education to estimate the wage gap between the private and the public sectors. The dataset represents the period 1993–2010, and is computed from three administrative registers: employment, education, and tax.[1] In addition to wages and education, we have information about the age, gender, immigrant status, industry affiliation, firm affiliation, and home region of all individuals. We concentrate on workers with full-time contracts (at least 30 hours per week).[2] Workers in the primary sectors of agriculture, fishing, and forestry are excluded, as well as workers below 22 years of age or above 65 years of age.

[1] The employment register links workers and firms and gives information on work contracts for all employees. It includes the length of the contract, and separates between full-time and part-time contracts. This is used to calculate the number of days worked per year, which is combined with data on annual wage income from the tax register to give a measure of daily wages. The education register covers the whole adult population and gives information about the highest completed education level in the beginning of October each year.

[2] Since the tax register gives information about total annual earnings, rather than separate earnings for each work contract, workers with more than two contracts during a year, as well as workers with one full-time and one part-time contract, are excluded. Workers with two full-time contracts are excluded if the number of days worked that year exceeds 455. This means that we allow for a maximum of three months of overlap between the two contracts. We also leave out workers with fewer than 89 working days during a year.

Finally, to avoid extreme observations, we exclude the top and bottom 1% of the wage distribution.

The original dataset consists of workers that remain in the private sector during the full period 1993–2010 ('private stayers'), workers that remain in the public sector ('public stayers'), and workers that change sector during the period ('shifters'). We start out with a comparison of shifters to the private sector with stayers in the public sector, while private stayers and shifters from the private to the public sector are excluded. We concentrate on public sector workers registered as public administration. Most public employees are registered in education and health care sub-sectors, but these sub-sectors are left out here – partly they are dominated by professions with particular characteristics, and partly they represent a mix of private and public sector workers. Since the wage formation for women is different from men, shown in a large literature dealing with gender, we concentrate on male workers. Further, the analysis considers high-educated workers since this is a more homogenous group. As shown by Rattsø and Stokke (2018), the low educated are more heterogeneous and face a more varied labor market.

Shifters to the private sector must be in the dataset for at least seven consecutive years: three years before the sector shift, the shift-year, and three years after the sector shift. Workers that are out of the labor market before the shift, as well as workers that shift back to the public sector within three years, are excluded. Sector shifts are observed during the years 1996 to 2007. We construct separate samples for the twelve shift years, each covering seven consecutive years. The analysis is based on the pooled samples, giving a panel from 1993 to 2010. Public stayers are included in a shift-year sample if observed for the full 7-year period. Among public stayers, the same worker can be part of several shift-year samples. As long as shifters do not change sector on January 1st, the wage in the shift-year represents a mix of private and public sector wages. We therefore exclude the shift-year, and compare the three years before the shift-year to the three years after the shift-year. The final dataset consists of 907,416 observations covering 25,361 different workers. Public sector stayers include 20,219 workers, while public-private shifters account for 5,142 workers.

Table 1 compares shifters and stayers with respect to observable worker characteristics. Shifters are younger than stayers (38.4 years of age compared to 44.1 years, measured in the shift year), and are more likely to live in a big city the year before the sector shift (defined as labor market regions with more than 150,000 inhabitants in 2010). We separate between two levels of higher education: postgraduate degree (more than four years) and some college education (1–4 years of duration). The share of workers with a postgraduate degree is 43% and 40% among shifters and stayers, respectively. When it comes to field of education, differences between the two groups of workers are larger. Shifters are much more likely to have an education in natural sciences, while stayers are overrepresented in educations related to transport, communications, and security.

Table 1 about here

As a further description of shifters versus stayers, we estimate the effect of worker characteristics on the probability of being a public-private shifter, documented in Table 2. The dependent variable is a dummy that equals one if the worker shifts from the public to the private sector during 1996–2007. Worker characteristics include age, level and field of education, and resident location. The regression includes shift year fixed effects. The findings supplement the descriptive statistics in Table 1. The impact of age on shifter probability is negative and convex. When a worker is one year older, the probability of being a shifter decreases by 0.4 percentage points (measured at 40 years of age). Having a postgraduate degree makes it 2 percentage points more likely to be a shifter, and living in a big city increases the probability by 0.7 percentage points. Compared to the reference category 'Humanities and arts', higher education in the field of 'Transport, communications, and security' makes it 3.8 percentage points less likely to be a public-private shifter, while an education in natural sciences increases the probability by 2.4 percentage points.

Table 2 about here

In the basic model, the estimation of wage effects from shifting to the private sector applies a difference-in-difference approach with shifters as the treatment group and public stayers as the control group. The analysis is based on variations of the following regression:

$$\ln w_{iarst} = \alpha \cdot shifter_i \cdot post\_shift_{it} + \lambda_a + \mu_r + \varphi_s + \gamma_t + \eta_i + \varepsilon_{iarst} \quad (1)$$

where w_{iarst} is the daily wage income for worker *i* in age group *a*, region *r*, shift year sample *s*, and year *t*, $shifter_i$ is a dummy that equals one if the worker has shifted from the public to the private sector during 1996–2007, and $post\_shift_{it}$ is a dummy that equals one in the three years following the sector shift. The regression controls for age group, regional, shift year, year and worker fixed effects, represented by λ_a, μ_r, φ_s, γ_t and η_i, respectively. We are interested in the interaction term between the shifter dummy and the dummy for post shift years. The parameter α captures the difference in wage change between treatment and control groups in the years after the sector shift compared to the pre-shift period. We also consider specifications with year-specific effects, where interaction terms between the shifter dummy and each year after the shift are included.

3. A basic difference-in-difference model

The difference-in-difference model described by equation (1) in section 2 allows for a comparison of shifters to the private sector (the treatment group) and stayers in the public sector (the control group) over time. The estimates show the wage effect of shifting from the public to the private sector. The identifying assumption is that the average development of wages among the stayers after the shift represents the counterfactual of the shifters. The model requires that treatment and control groups follow parallel trends in daily wages during the pre-shift years. Figure 1 shows the wage time trends of stayers versus shifters three years before and after the shift-year. To enable comparison across different shift years, we focus on the development of a daily wage index set to equal 100 three years before the sector shift (denoted t–3 in the figure). In the years before the shift, the two groups of workers have roughly similar wage trends, but the wages of shifters increase somewhat in year t–1 compared to stayers. The difference in pre-trend indicates that the identifying assumption is rejected (see below). In the shift year, the wage path of shifters increases rapidly and the

difference in wage paths between shifters and stayers continues to increase in the years following the shift.

Figure 1 about here

The difference-in-difference model is given three specifications: the first studying the average post-shift wage of the shifters, the second estimating the effect of the three years after the shift separately, and the third including also the pre-shift year effect. The approach addresses workers with common background in the public sector, and the analysis compares public and private sector wages taking into account unobservable worker characteristics, regional differences, and differences between age groups. The estimation results are given in Table 3.

High-educated male workers have a gain by shifting to the private sector of 11.6% for the three-year period following the shift, as seen from column 1. Separating out the effect for each year in column 2, we see that the wage gain is increasing from 8.8% in the first year after the shift to 12.5% and 13.7% in the second and third year after the shift, respectively. The increasing premium reflects the higher return to experience in the private sector, consistent with the findings of Rattsø and Stokke (2018). The more flexible dynamic specification in column 3 tests for the wage effect for shifters before the shift year by including interaction terms between the shifter dummy and pre-shift years. The identifying assumption is that these interaction terms are not statistically significant, confirming parallel wage trends between shifters and stayers in the years before the shift. Consistent with the pattern in Figure 1, the wage effect in the year before the shift is significantly higher for shifters than for stayers, although the size of the effect is small. We conclude that shifters have different wage development from stayers before the shift. The estimate of the model cannot be interpreted as a causal wage effect of the shift.

The size of the wage gap between private and public sectors vary with the wage institutions of the country. Existing studies of other countries estimate gaps for high-educated of about the same size. Schanzenbach (2015) estimate a 'public sector pay penalty' of about 9% for workers with college degree (excluding teachers) in the US. Controlling for college major categories and occupation, the static wage gap is reduced to 5-6%.

Table 3 about here

To investigate the role of unobservables, we estimate the model without worker fixed effects in Appendix Table 1. Compared with the post-shift effect of 11.6%, we now get 12.6%. Excluding worker fixed effects, overestimates the return to shifting to the private sector by 1 percentage point, or almost 10%. It follows that the shifters to the private sector are positively selected – they have higher abilities than the stayers.

4. Identification based on early vs. late shifters

The basic difference-in-difference model in Table 3 narrows down the comparison of treatment and control groups to individuals with common background in the public sector, but the analysis rejects the identifying assumption of this model. The wage trends of stayers and shifters are not parallel in the years before the sector shift, and consequently, public stayers are not a valid control group for public-private shifters. We suggest an identification strategy concentrating on the shifters. The argument is that shifters are different from stayers and have some common characteristics relevant for the wage formation.

Our treatment group consists of workers that shift from the public to the private sector during 1996–2001, referred to as early shifters. Early shifters are included in the dataset three years before and after the shift year and are therefore observed during 1993–2004. The control group consists of workers that shift to the private sector in 2002 or later, referred to as late shifters. We only include observations of late shifters in the years before they shift sector (while still working in the public sector). The dataset is constructed in the same manner as in section 2, with separate samples for each of the six shift-years pooled together to a panel covering 1993–2004. The dataset consists of 43,343 observations and 4,303 different workers. Early shifters account for 3,098 workers, while late shifters account for 1,205 workers. Both early and late shifters are observed three years before and after the shift-year of early shifters (which explains why some late shifters are excluded from the control group).

The underlying assumption of identification is that early and late shifters are similar with respect to unobserved characteristics important for the wage effect of shifting to the private sector. To get a measure of the possible bias related to selection, we compare the private sector wage premium estimated from an analysis of early vs. late shifters to the basic difference-in-difference model with public sector stayers as the control group. The findings are given in Table 4. The treatment group consists of early shifters. In columns (1) – (3), the control group is restricted to stayers covering the period of early shifters (1993–2004), referred to as early stayers. Columns (4) – (6) represent our preferred specification, where late shifters (in the years before they shift) are the control group. The main result in terms of identification is that the parallel paths assumption holds for the shifter model. The pre-shift difference in wage paths is reproduced when stayers are the control group (column 3), but disappears when they are replaced by late shifters (column 6). There are parallel trends in wages before the shift for early and late shifters. The identifying assumption for a causal effect of shifting now holds. This is confirmed by Figure 2, which illustrates the development of the daily wage index for early and late shifters three years before and after the sector shift for early shifters. The wage trend in the years before the sector shift is similar for the two groups of workers, and comparing Figures 1 and 2 confirms that shifters are a more homogeneous group than shifters and stayers.

Figure 2 about here

Comparing columns (1) and (4) of Table 4, the estimated average wage gain is reduced from 12% to 10% when late shifters (rather than early stayers) are used as control group. The reduction of the estimated wage effect from the basic model to the shifter model measures a possible bias when not taking into account the selection of stayers into shifters. The selection problem implies overestimation of the wage gain from shifting to the private sector. The bias is 2 percentage points or about 20%. The overestimation of the private-public wage premium represents a positive selection of shifters to the private sector compared to stayers in the public sector. When we concentrate on shifters only, we correct for this source of bias. Columns (2) and (3) versus (5) and (6) in Table 4 investigate the dynamics of adjustment in the basic difference-in-difference model versus the shifter model. The private wage premium is increasing over time in both models. Our understanding is that there is higher return to

experience accumulated in the private sector.

Table 4 about here

There is a concern that early and late shifters are different and shift to the private sector for different reasons. To follow up on this, Table 5 compares observable characteristics of early shifters, late shifters, and early stayers. The mean values of age refer to the shift-year of early shifters and resident location is based on the year before the sector shift for early shifters. Other variables are constant over time. While early shifters are younger than early stayers (consistent with the descriptive statistics of all shifters and stayers in Table 1), both early and late shifters are on average 38 years of age at the time when early shifters change sector. More importantly, the age distributions of early and late shifters (measured in the shift-year of early shifters) have the same shape and differ significantly from the age distribution of early stayers, as documented in Figure 3. In addition, early and late shifters are more comparable when it comes to field of education and likelihood of living in a big city. These descriptive statistics give support to our interpretation of the shifter analysis as a measure of bias in the basic model.

Table 5 and Figure 3 about here

As seen from Table 5, early and late shifters are on average at the same age when early shifters change sector. This implies that late shifters on average shift sector somewhat later in the career compared to early shifters. While shifters during 1996–2001 are on average 38 years of age in the shift-year, shifters during 2002–2007 are on average 39.2 years of age when they shift sector. In addition, our data restriction to include only those late shifters who are observed for the full 7-year period before their own shift, contributes to a sample of late shifters that are at the same age as early shifters when the latter group shift sector. As a check of robustness, we also do the analysis based on all late shifters, independent of how many years they are observed in the data. When considering all late shifters, the average age in the shift-year of early shifters is 36 years. The findings remain the same, as documented in Appendix Table 2.

5. Identification based on business cycles

As a robustness check, we suggest an alternative identification strategy that takes benefit of different recruitment to the public sector in recessions and booms. The individuals recruited to the public sector in booms versus recessions are selected in different ways, and the boom-recruited are those preferring to work in the public sector. Introduction of separate shift effects for individuals recruited in boom and recession allows for a measure of the selection bias. The shifters recruited during recession are expected to have higher return because they are different from those staying in the public sector.

During the period studied, the public sector has been steadily increasing financed by rising government oil revenues. The private sector, on the other hand, has experienced recessions and booms, in particular the 'bank crisis' recession in the early 1990s (1993–1996 in our data), followed by boom 1997–2001, and then recession 2002–2005 driven by international contraction. We concentrate on the economic conditions in the year of labor market entry. In boom years, the desire for jobs in the public sector is limited to individuals with special preferences/motivation. Other people seek private sector employment in booms. In recessions, many individuals ready for work in the private sector are 'forced' to enter the public sector. It follows that the individuals recruited to the public sector in booms versus recessions are selected in different ways. Using a difference-in-difference-in-difference model, we can compare the returns of shifting to the private sector for individuals recruited in booms and recessions.

The analysis is restricted to shifters who enter the labor market in 1993 or later, which implies young workers born after 1967. Among the 5,142 high-educated shifters in the basic model of Table 3, we can identify the year of labor market entry for 1,315 shifters (452 entering in boom years and 863 entering in recession years). We include public stayers born after 1967, and in total, the data cover 6,571 workers. The basic difference-in-difference model for this sample of workers gives results comparable to Table 3 (not reported here). The introduction of separate shift effects for individuals recruited in boom years (reference case) and recession years (interaction terms) shows significant differences between the two groups, as

documented in Table 6. In column 1, the average effect of the three years following the shift is 9.4% for those recruited in boom years and an additional 2.9% for those recruited in recession years (significant at the 10% level). Column 2 reports estimates separating the shift-effect for the three years following the shift. The difference in wage premium is mainly related to the first year after the shift, where the wage gain is 3.6 percentage points higher for the recession-recruited (significant at the 5% level).

Table 6 about here

Our understanding is that workers recruited to the public sector during booms represent those with preferences to work in the public sector. They are consequently most similar to the stayers. Those recruited during recessions are different, they would like to work in the private sector, and they gain more by shifting to the private sector. The difference in return between those recruited in boom and recession is a measure of the bias you get because shifters are different from the stayers. Based on the average effect for the three years following the shift (given in column 1), the recession-recruited have about 30% more return from shifting to the private sector compared to the boom-recruited.

The main assumption for the estimation of the bias is that there is important variation in abilities and/or motivations to shift from the public to the private sector among those recruited in different business cycle situations. Those recruited to the public sector during recession are more likely to shift to the private sector than those recruited during boom. The last group is more similar to the stayers in the public sector, and the difference in private wage premium between the two groups reflects selection. The main concern is that we do not have two distinct types of individuals, and that there is heterogeneity in abilities and motivations in both groups. The significant difference in return to shift between those recruited in recession versus boom indicates that the business cycles matter for recruitment to the public sector. The analysis does not address the reasons for shifting to the private sector. Future research will look into this heterogeneity in terms of occupations and firms in the private sector.

6. Concluding remarks

The wage gap between the private and public sectors is analyzed in a difference-in-difference model, and in the basic version, we compare shifters from the public to the private sector with stayers in the public sector. In an extension of the basic model, we suggest an identification strategy comparing early shifters with workers still in the public sector, but shifting later. The model satisfies the identifying assumption of the difference-in-difference model with parallel wage paths in the years before the sector shift. The analysis implies an overestimation of the wage gap by 20% in the difference-in-difference model comparing shifters with stayers. The overestimation bias represents a positive selection of shifters to the private sector compared to stayers in the public sector. An alternative identification strategy based on different recruitment to the public sector in booms versus recessions, confirms the overestimation in the basic model. It should be noticed that further methodological challenges related to reverse causality remain.

References

Bargain, O. and B. Melly (2008). Public sector pay gap in France: New evidence using panel data. IZA Discussion Paper 3427.

Bargain, O., A. Etienne and B. Melly (2016). The public sector wage gap: New evidence from panel administrative data. Mimeo, Aix-Marseille University.

Borjas, G. (2003). The wage structure and the sorting of workers into the public sector. In J.D. Donahue and J.S. Nye (eds.), *For the People: Can We Fix Public Service*, Brookings Institution Press.

Depalo, D. (2017). Identification issues in the public/private wage gap, with an application to Italy. *Journal of Applied Econometrics* 33, 435-456.

Dickson, M., F. Postel-Vinay, and H. Turon (2014). The lifetime earnings premium in the public sector: The view from Europe. *Labour Economics* 31, 141-161.

Giordano, R., M. Pereira, D. Depalo, B. Eugene, E. Papapetrou, J. Perez, L. Reiss, and M. Roter (2014). The public sector pay gap in a selection of euro area countries in the pre-crisis period. *Hacienda Publica Espanola* 214, 3, 11-34.

Hospido, L. and E. Noral-Benito (2016). The public sector wage premium in Spain: Evidence from Longitudinal administrative data, *Labour Economics* 42, 101-122.

Lausev, J. (2014). What has 20 years of public-private pay gap literature told us? Eastern European transitioning vs. developed economies. *Journal of Economic Surveys* 28, 3, 516-550.

Postel-Vinay, F. and H. Turon (2007). The public pay gap in Britain: Small differences that (don't?) matter. *Economic Journal* 117, 1460-1503.

Rattsø, J. and H.E. Stokke (2018). Dynamic private-public wage gap: Estimates based on sector-specific experience data. Mimeo, Department of Economics, Norwegian University of Science and Technology.

Schanzenbach, M. (2015). Explaining the public-sector pay gap: The role of skill and college major. *Journal of Human Capital* 9, 1, 1-44.

Figure 1

Wage trend, shifters vs. stayers

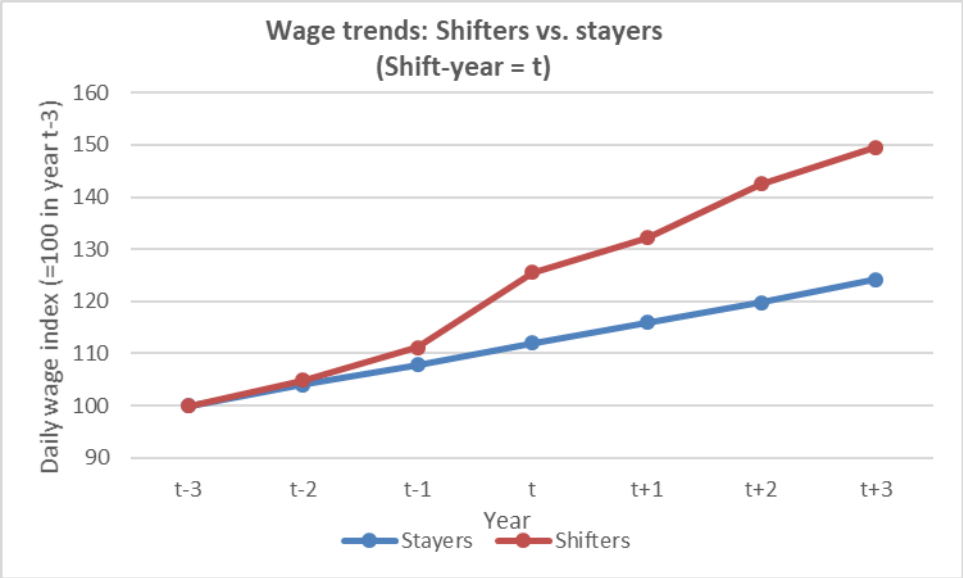

Figure 2

Wage trend, early shifters vs. late shifters (before shifting)

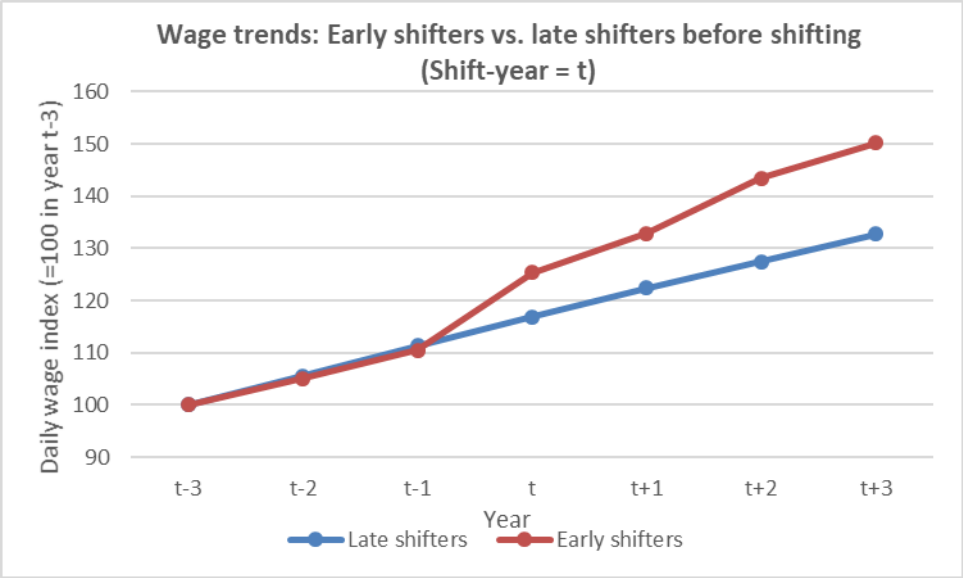

Figure 3

Age distribution of early shifters, late shifters, and early stayers (measured in the shift-year of early shifters)

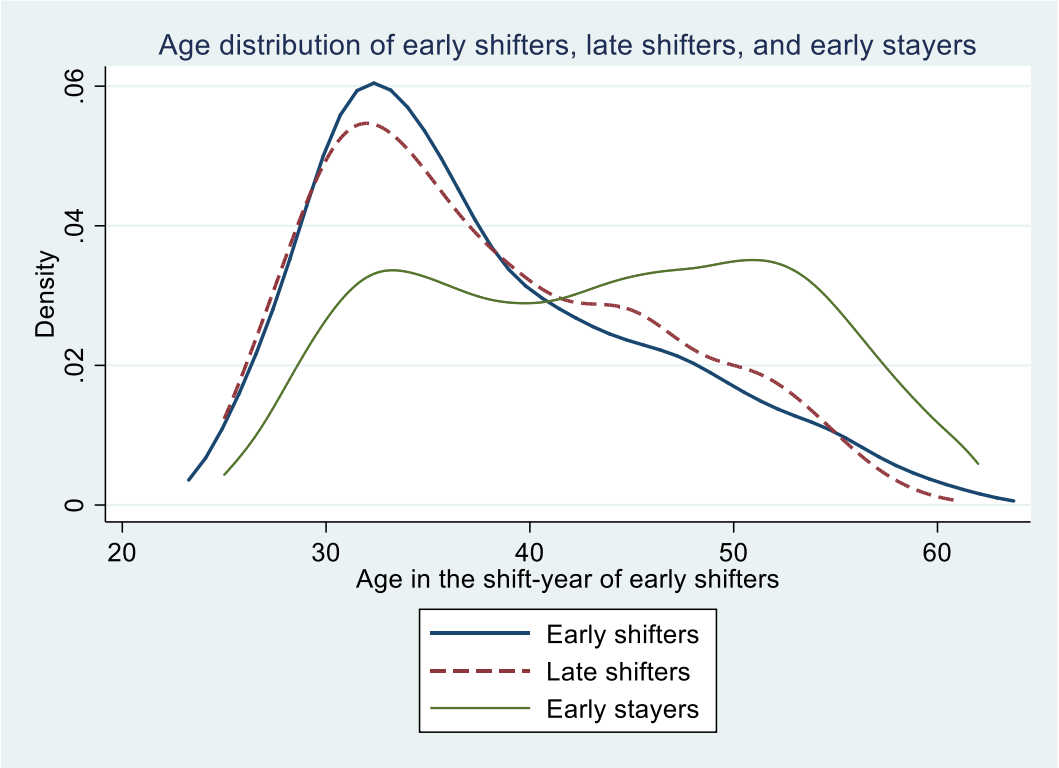

Table 1

Descriptive statistics: Public-private shifters vs. public stayers (mean values)

	Public-private shifters	Public stayers
Age	38.4	44.1
Some college education	0.572	0.604
Postgraduate degree	0.428	0.396
Field of education		
Humanities and arts	0.042	0.052
Teacher training and pedagogy	0.033	0.05
Social sciences and law	0.175	0.19
Business and administration	0.17	0.15
Natural sciences	0.347	0.22
Health, welfare and sport	0.012	0.017
Primary industries	0.024	0.029
Transport, communications, and security	0.188	0.285
Big city resident	0.52	0.448
Small city resident	0.206	0.227
No. of workers	5,142	20,219

Notes: We separate between two levels of higher education: 'Some college education' is defined as 1 – 4 years at university or college, while a postgraduate degree has a duration of more than four years. Three region types: Big cities (more than 150,000 inhabitants in 2010) accounting for 7 out of 89 labor market regions, small cities (population in the range 65,000 – 150,000 in 2010, 13 regions), and the remaining 69 regions. The mean value of age refers to the shift year, while the resident location is based on the year before the sector shift. Other variables are constant over time.

Table 2

Probability of being a public-private shifter (vs. public stayer)

	(1)
Dependent variable	Shifter
Age	-0.012***
	(0.0005)
Age^2	0.0001***
	(0.0000)
Postgraduate degree	0.02***
	(0.0011)
Big city resident	0.007***
	(0.0011)
Small city resident	0.004***
	(0.0013)
Teacher training and pedagogy	0.012***
	(0.0029)
Social sciences and law	-0.008***
	(0.0024)
Business and administration	0.01***
	(0.0024)
Natural sciences	0.024***
	(0.0023)
Health, welfare and sport	0.005
	(0.0041)
Primary industries	-0.006
	(0.0035)
Transport, communications, and security	-0.038***
	(0.0023)
Shift year fixed effects	Yes
Obs.	151,236
No. of workers	25,361
No. of public stayers	20,219
No. of shifters	5,142
R^2	0.03

Notes: The dependent variable is a dummy that equals one if the worker shifts from the public to the private sector during 1996-2007. Resident location is measured the year before the sector shift, while other variables are measured in the shift year. The reference category for field of education is 'Humanities and arts'. *** indicates significance at the 1 percent level. The regression includes a constant term.

Table 3

Basic difference-in-difference model with public stayers as the control group

Dependent variable	(1) Log daily wage	(2) Log daily wage	(3) Log daily wage
Shifter x post shift-year	0.116***		
	(0.0034)		
Shifter x shift-year$_{t-2}$			0.003
			(0.0021)
Shifter x shift-year$_{t-1}$			0.013***
			(0.0027)
Shifter x shift-year$_{t+1}$		0.088***	0.093***
		(0.0034)	(0.0038)
Shifter x shift-year$_{t+2}$		0.125***	0.131***
		(0.0037)	(0.004)
Shifter x shift-year$_{t+3}$		0.137***	0.143***
		(0.004)	(0.0043)
Year fixed effects	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes
Shift year fixed effects	Yes	Yes	Yes
Age group fixed effects	Yes	Yes	Yes
Obs.	907,416	907,416	907,416
No. of workers	25,361	25,361	25,361
No. of public stayers	20,219	20,219	20,219
No. of shifters	5,142	5,142	5,142
R^2	0.81	0.81	0.81

Notes: The regressions are based on yearly data during 1993-2010 for public stayers and public-private shifters. The dependent variable is log daily wages in constant 2010 prices. Regional fixed effects include 89 labor market regions, and shift year fixed effects refer to twelve separate shift-year samples during 1996-2007. The age controls are given as five-year intervals. Robust standard errors (clustered by workers) are given in parenthesis. *** indicates significance at the 1 percent level. All regressions include a constant term.

Table 4

Identification based on early vs. late shifters

	Early stayers as control			Late shifters as control		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Log daily wage	Log daily wage	Log daily wage	Log daily wage	Log daily wage	Log daily wage
Shifter x post shift-year	0.12***			0.098***		
	(0.0044)			(0.0057)		
Shifter x shift-year$_{t-2}$			0.003			-0.002
			(0.0026)			(0.0031)
Shifter x shift-year$_{t-1}$			0.008**			-0.002
			(0.0033)			(0.0042)
Shifter x shift-year$_{t+1}$		0.093***	0.096***		0.076***	0.075***
		(0.0043)	(0.0048)		(0.0054)	(0.0064)
Shifter x shift-year$_{t+2}$		0.129***	0.132***		0.11***	0.108***
		(0.0047)	(0.0051)		(0.0061)	(0.0071)
Shifter x shift-year$_{t+3}$		0.14***	0.144***		0.119***	0.118***
		(0.0052)	(0.0056)		(0.0069)	(0.0078)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Shift year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Age group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	433,686	433,686	433,686	43,343	43,343	43,343
No. of workers	18,648	18,648	18,648	4,303	4,303	4,303
No. of early stayers	15,550	15,550	15,550			
No. of late shifters				1,205	1,205	1,205
No. of early shifters	3,098	3,098	3,098	3,098	3,098	3,098
R^2	0.80	0.80	0.80	0.75	0.75	0.75

Notes: The treatment group consists of workers that shift from the public to the private sector during 1996 – 2001, referred to as early shifters. In columns (1) – (3), the control group consists of public sector stayers during the early period (1993 – 2004). In columns (4) – (6), the control group consists of workers that shift to the private sector in 2002 or later, referred to as late shifters. We only include observations of late shifters in the years before they shift sector (while still working in the public sector). Early and late shifters, as well as early stayers, are in the dataset three years before/after the shift of early shifters and are therefore observed during 1993 – 2004. The dependent variable is log daily wages in constant 2010 prices. Robust standard errors (clustered by workers) are given in parenthesis. *** and ** indicate significance at the 1 and 5 percent levels, respectively. All regressions include a constant term.

Table 5

Descriptive statistics: Early shifters, early stayers, and late shifters before shifting (mean values)

	Early shifters	Early stayers	Late shifters before shifting
Age	37.9	43.4	38.0
Some college education	0.604	0.588	0.595
Postgraduate degree	0.396	0.412	0.405
Field of education			
Humanities and arts	0.04	0.054	0.042
Teacher training and pedagogy	0.036	0.054	0.048
Social sciences and law	0.149	0.187	0.152
Business and administration	0.17	0.15	0.135
Natural sciences	0.358	0.233	0.276
Health, welfare and sport	0.013	0.018	0.02
Primary industries	0.02	0.029	0.029
Transport, communications, and security	0.208	0.268	0.286
Big city resident	0.49	0.442	0.471
Small city resident	0.214	0.225	0.221
No. of workers	3,098	15,550	1,205

Notes: We define early shifters as workers that shift from the public to the private sector during 1996 – 2001, and early stayers as workers that are in the public sector during the early period and never shift to the private sector. The group of late shifters consists of workers that are in the public sector in the early period, but who shift to the private sector in 2002 or later. Variables are described in the notes to Table 1. The mean value of age refers to the shift year of early shifters, while the resident location is based on the year before the sector shift for early shifters. Other variables are constant over time.

Table 6

Identification based on business cycles at labor market entry

Dependent variable	(1) Log daily wage	(2) Log daily wage	(3) Log daily wage
Shifter x post shift-year	0.094***		
	(0.0126)		
Shifter x shift-year$_{t-2}$			0.014
			(0.0092)
Shifter x shift-year$_{t-1}$			0.038***
			(0.0116)
Shifter x shift-year$_{t+1}$		0.05***	0.067***
		(0.0125)	(0.0142)
Shifter x shift-year$_{t+2}$		0.111***	0.128***
		(0.0134)	(0.015)
Shifter x shift-year$_{t+3}$		0.122***	0.139***
		(0.015)	(0.0165)
Shifter x post shift-year x recession entry	0.029*		
	(0.0157)		
Shifter x shift-year$_{t-2}$ x recession entry			-0.01
			(0.011)
Shifter x shift-year$_{t-1}$ x recession entry			-0.019
			(0.014)
Shifter x shift-year$_{t+1}$ x recession entry		0.036**	0.027
		(0.0157)	(0.0177)
Shifter x shift-year$_{t+2}$ x recession entry		0.027	0.018
		(0.0167)	(0.0186)
Shifter x shift-year$_{t+3}$ x recession entry		0.022	0.012
		(0.0185)	(0.0201)
Year fixed effects	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes
Shift year fixed effects	Yes	Yes	Yes
Age group fixed effects	Yes	Yes	Yes
Obs.	171,072	171,072	171,072
No. of workers	6,571	6,571	6,571
No. of public stayers	5,256	5,256	5,256
No. of shifters	1,315	1,315	1,315
No. of shifters entering in boom	452	452	452
No. of shifters entering in recession	863	863	863
R^2	0.73	0.73	0.73

Notes: The regressions are based on yearly data during 1993-2010 for public stayers and public-private shifters. The analysis is restricted to shifters who enter the labor market in 1993 or later, which implies young workers born after 1967. The control group consists of public stayers born after 1967. We allow for separate shift effects for individuals recruited in boom years (reference case) and recession years (interaction terms). The dependent variable is log daily wages in constant 2010 prices. Robust standard errors (clustered by workers) are given in parenthesis. ***, ** and * indicate significance at the 1, 5 and 10 percent level, respectively. All regressions include a constant term.

Appendix Table 1
Basic difference-in-difference model without worker fixed effects

Dependent variable	(1) Log daily wage	(2) Log daily wage	(3) Log daily wage
Shifter	-0.032***	-0.032***	-0.041***
	(0.0032)	(0.0032)	(0.0035)
Shifter x post shift-year	0.126***		
	(0.0034)		
Shifter x shift-year$_{t-2}$			0.006***
			(0.002)
Shifter x shift-year$_{t-1}$			0.021***
			(0.0027)
Shifter x shift-year$_{t+1}$		0.096***	0.105***
		(0.0034)	(0.0038)
Shifter x shift-year$_{t+2}$		0.135***	0.144***
		(0.0036)	(0.004)
Shifter x shift-year$_{t+3}$		0.149***	0.158***
		(0.004)	(0.0044)
Worker characteristics	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes
Worker fixed effects	No	No	No
Shift year fixed effects	Yes	Yes	Yes
Age group fixed effects	Yes	Yes	Yes
Obs.	907,416	907,416	907,416
No. of workers	25,361	25,361	25,361
No. of public stayers	20,219	20,219	20,219
No. of shifters	5,142	5,142	5,142
R^2	0.30	0.30	0.30

Notes: The regressions are based on yearly data during 1993-2010 for public stayers and public-private shifters. The dependent variable is log daily wages in constant 2010 prices. Worker characteristics include immigrant status and level of education (within the broader group of higher education). Robust standard errors (clustered by workers) are given in parenthesis. *** indicates significance at the 1 percent level. All regressions include a constant term.

Appendix Table 2

Identification based on early vs. late shifters: Less restrictive definition of late shifters

	Late shifters as control		
	(1)	(2)	(3)
Dependent variable	Log daily wage	Log daily wage	Log daily wage
Shifter x post shift-year	0.1***		
	(0.0055)		
Shifter x shift-year$_{t-2}$			0.000
			(0.003)
Shifter x shift-year$_{t-1}$			0.002
			(0.0041)
Shifter x shift-year$_{t+1}$		0.077***	0.078***
		(0.0052)	(0.0062)
Shifter x shift-year$_{t+2}$		0.11***	0.111***
		(0.0058)	(0.0067)
Shifter x shift-year$_{t+3}$		0.12***	0.121***
		(0.0065)	(0.0073)
Year fixed effects	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes
Shift year fixed effects	Yes	Yes	Yes
Age group fixed effects	Yes	Yes	Yes
Obs.	82,030	82,030	82,030
No. of workers	6,924	6,924	6,924
No. of late shifters	3,826	3,826	3,826
No. of early shifters	3,098	3,098	3,098
R^2	0.73	0.73	0.73

Notes: The treatment group consists of workers that shift from the public to the private sector during 1996 – 2001, referred to as early shifters. Early shifters are included in the dataset three years before/after the shift and are therefore observed during 1993 – 2004. The control group consists of workers that shift to the private sector in 2002 or later, referred to as late shifters. We only include observations of late shifters in the years before they shift sector (while still working in the public sector). Late shifters are included even though they are not observed during the full 7-year period around the shift of the early shifter. The dependent variable is log daily wages in constant 2010 prices. Robust standard errors (clustered by workers) are given in parenthesis. *** and ** indicate significance at the 1 and 5 percent levels, respectively. All regressions include a constant term.

ISSN 1503-299X

WORKING PAPER SERIES

No. 1/2018

Identification of the private-public wage gap*

Jørn Rattsø
Department of Economics
Norwegian University of Science and Technology

Hildegunn Ekroll Stokke
Department of Economics
Norwegian University of Science and Technology

Department of Economics

Norwegian University of Science and Technology
N-7491 Trondheim, Norway
http://www.ntnu.edu/econ/working-papers

* We appreciate discussions at the 12th winter school on inequality and social welfare theory in Canazei, the 2017 European Meeting of the Urban Economics Association in Copenhagen, the 2017 LAGV conference in Aix, the 2018 DEMOS-workshop on urban/labor economics, and the 2018 EALE conference in Lyon, and in particular comments from David Albouy, Nate Baum-Snow, Paolo Naticchioni, Kjell G. Salvanes, and Olmo Silva. We are grateful for the cooperation of Statistics Norway and funding from the Research Council of Norway (grant number 255509).
** Corresponding author.