

Maczulskij, Terhi

**Conference Paper**

## Employment sector and pay gaps: genetic and environmental influences

52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Maczulskij, Terhi (2012) : Employment sector and pay gaps: genetic and environmental influences, 52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/120683>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Employment sector and pay gaps: genetic and environmental influences<sup>\*</sup>

Terhi Maczulskij

University of Jyväskylä, School of Business and Economics, P.O. Box 35, FI-40100, Jyväskylä, Finland

Email: terhi.a1.maczulskij@jyu.fi, Phone: +358405767824, Fax: +358(0)14617194

## Abstract

This article investigates the influence of genetic factors and shared environments on public sector employment choice and public-private sector pay gaps. The analyses use data on Finnish twins for the period from 1990 to 2004. The results indicate that approximately 30-40 per cent of the observed variance in public sector employment is attributed to genetic factors and that the role of shared environment is statistically zero. Additional findings suggest that genetic factors may encourage people to engage in public sector employment through at least one mechanism: educational attainment. The pay gap analysis provides evidence that females receive 4-5 per cent less in the public sector and indicates that this gap is explained by neither genetic effects nor the influence of a shared environment. For males, the pay gap is zero rather than negative 7 per cent as conventionally estimated, and this difference is mainly caused by genetic factors such as intelligence.

**Keywords:** public sector employment; behavioural genetics; twin studies; wage differentials

**JEL classifications:** J24; J31; J45

---

<sup>\*</sup> This paper is part of the research project “Research on entrepreneurship, labour market outcomes and education using twin data” funded by the Academy of Finland (number 127796). I thank Ari Hyytinen, Otto Toivanen, Pekka Ilmakunnas, Jaakko Pehkonen, Merja Kauhanen, Jaana Rahko, Petri Böckerman, all of the members of the research project and the participants at the Allecon seminar in Tampere (December 2010), at the Summer seminar in economics in Jyväskylä (June 2011 and 2012) and at the XXXIV Annual Meeting of the Finnish Economy Association in Vaasa (February 2012) for providing helpful comments.

# **1. Introduction**

Internationally and also in Finland, there has been an interest in public sector labour markets for several decades. Traditionally, research on this field has addressed two main questions. The first is that what types of individuals are attracted to public sector jobs and what attributes makes those jobs appealing, and the second relates to the differences between the pay earned by public and private sector employees.

It is standard to consider public sector employment as an endogenous outcome variable. The literature has shown that a number of demographic factors contribute positively to the decision to become a public sector worker. These include higher age and education<sup>1</sup>, being married, being from an ethnic minority group, being a house renter and being raised in a family in which a parent or other relative worked for the government (e.g., Bellante and Link 1981, Ophem 1993, Kanellopoulos 1997, Dustmann et al. 1998, Adamchick and Bedi 2000, Christofides and Pashardes 2002, Lewis and Frank 2002). The second group of attributes relates to individual psychological factors. Many studies have found that some individuals are more attracted to and motivated to begin public service work (e.g., Perry and Wise 1990); public sector employees value non-monetary benefits more highly (e.g., Crewson 1997, Karl and Sutton 1998), and public sector jobs are especially attractive to risk-averse individuals

---

<sup>1</sup> Using data for German males, Dustmann and van Soest (1998) found that the positive effect of education on public sector employment choice disappears after controlling education with family background variables. Based on this finding, they conclude that the positive relationship found in standard models is not structural but rather reflects unobserved heterogeneity.

(Bellante et al. 1981, Christofides et al. 2002, Hartog et al. 2002, Pfeifer 2011).<sup>2</sup> In addition, the public sector has long experienced serious difficulty recruiting high-quality labour. In the UK, Nickell and Quintini (2005) found that for male employees, decreasing relative pay in the public sector resulted in a measurable decrease in quality. The same phenomenon did not appear to be applicable to female workers, although evidence of the reduced ability of female teachers is commonly reported in the US (e.g., Corcoran et al. 2004).

The second commonly discussed issue regarding public sector labour markets is the wage differential between public and private sector employees. The literature provides varying empirical evidence regarding the existence of pay gaps in the labour market. These findings indicate that employees typically earn less in the private sector than in the public sector. However, these results are not commented on here because they are not of direct interest. Instead, we will examine a more important question: why such pay gaps exist. Kanellopoulos (1997), Lassibille (1998) and Christofides and Pashardes (2002) suggest that these pay gaps are mainly explained by differences in individual attributes such as better skill characteristics. The effect of qualifications on

---

<sup>2</sup> Bellante et al. (1981) examined risk aversion using an index formulated on the basis of answers to questions involving the condition and insurance of automobile owners, the use of seat belts, medical coverage and smoking and drinking habits. Christofides et al. (2002) used life insurance and loan repayments as proxies for attitudes regarding risk. Hartog et al. (2002) measured risk aversion by asking individuals to state their reservation price for a lottery ticket after specifying the probability of winning a prize of a particular magnitude. Using expected utility theory, Hartog et al. then deduced the Arrow-Pratt measure for risk aversion. In his main analysis, Pfeifer (2011) used a 2004 wave from a longitudinal survey of private households and persons in Germany that included questions about participants' risk-taking behaviour in general and their careers in particular.

pay gaps is typically larger for males, whereas females are sometimes seen as receiving a pure pay advantage. Bender and Elliot (2002) extend the conventional analysis by emphasising the role of job attributes in the unequal pay offered by the two sectors. They find that in addition to human capital characteristics, job-specific attributes, such as the use of computers and establishment size, account for a major part of the observed pay gap between public and private sector employees.<sup>3</sup> Lucifora and Meurs (2006) present another view of the existence of pay gaps on a macro level. Using data for the UK, France and Italy, they observe that the gaps are actually larger in countries in which pay formation is less regulated.

However, our current knowledge about public sector labour markets is far from being complete, and many aspects of the two questions still remain unanswered. To the best of my knowledge, no study has attempted to determine the role of genetics in public sector employment choice and pay gaps. This is relatively surprising given that the behavioral genetics literature provides convincing empirical evidence regarding heritable effects on a variety of sociological and economic outcomes. This branch of the literature includes Miller et al. (2001), who investigate the genetic heritability of education; Johnson et al. (2004), who find that marital status is mainly explained by genetic factors; Whiston and Keller (2004) for their review of occupational preferences; Keller et al. (1992), who study the heritability of work values; Plomin and Spinath (2004), who focus on ability; and Cesarini et al. (2009), who examine the genetic heritability of risk-taking. Because all of these attributes also contribute to public sector

---

<sup>3</sup> See also Maczulskij and Pehkonen (2011), who found evidence that differences in industry-specific valuations of skills explain a significant part of the pay gaps in Finland (together with human capital variables).

employment choice, it is reasonable to ask to what extent public sector employment is in fact explained by genetic factors. In addition to genetic effects, public sector employment may also be influenced by shared environments, which may include role-modelling and the transmission of labour market information from parents to children along the lines of Dustmann et al. (1998) and Lewis et al. (2002).

As has been well established, family background, personal traits, intelligence and other unobserved characteristics affect individual earnings. A particular challenge in the empirical literature is the fact that the object of primary interest, ability, can often be measured only using poor proxies. Therefore, the results obtained from the traditional earnings research and those obtained using pay gap estimation may be seriously biased if ability is not controlled for in the analysis.

The goal of this study is to fill the gap in the literature by investigating the impacts of shared environment and genetic factors on public sector employment choice and public-private sector pay gaps. This is accomplished by using data on twins as twin siblings have a similar family background and identical twins are also similar in terms of their genetics. The results of this article suggest that genetic factors indeed account for a large part of the observed variance in public sector employment (approximately 30-40 per cent) and that the effect of shared environments is statistically insignificant. The pay gap analyses indicate that the true pay gap for males is zero rather than negative 7 per cent as conventionally estimated and that this bias is mainly attributed to genetic factors such as intelligence and risk-taking behaviour. The same phenomenon does not appear important for females, who nevertheless receive at least an economically significant pay disadvantage of 4-5 per cent in the public sector.

The rest of this paper is organised as follows. Section 2 presents the twin sample, which is matched with the Finnish Longitudinal Employer-Employee Data (FLEED). Section 3 presents the empirical analysis of the effects of shared environment and genetic factors on public sector employment choice. Section 4 reports similar findings based on wage differential analysis. Finally, Section 5 concludes the paper.

## **2. Data**

### **2.1 Data sources**

The twin sample is based on the Finnish Twin Cohort Study by the Department of Public Health at the University of Helsinki, which has been matched to the Finnish Longitudinal Employer-Employee Data (FLEED) from Statistics Finland.

The Twin Cohort Study was established in 1974 and was initially conducted using the Central Population Registry of Finland. The original study involved the use of a postal health survey that was conducted in 1975, 1981 and 1990 on same-gender twin pairs who were born before 1958. The third questionnaire was sent to pairs who were born between the years 1930 and 1957 and who were thus between the ages of 33 and 60 in 1990. A total of 16,179 twin pairs were contacted, and the response rate for the 1990 survey was 77 per cent. The final number of twin pairs in the sample is 12,502, giving us 25,004 individuals in total.

The twin pairs are either fraternal (dizygotic, DZ) or identical (monozygotic, MZ).<sup>4</sup> The zygosity of the twin pairs was determined based on their answers to questions regarding their similarity in appearance during childhood. A subsample was reclassified using 11 blood markers. The classification results were completely consistent, with a 1.7 per cent probability of the misclassification of a blood marker-concordant pair (Kaprio et al. 1979). The survey contains information on symptoms of illnesses and reported diseases, drug use, physical characteristics, smoking, alcohol use, leisure time physical activity and psycho-social factors.

The original FLEED cover nearly all Finns between the ages of 16 and 70 for the period from 1988 to 2004. The data are constructed from a number of different administrative registers for individuals, firms and establishments that are collected or maintained by Statistics Finland. Using a personal identifier, the Twin Cohort Study is matched with the FLEED. The final sample contains information for the period from 1990 to 2004 on 12,502 twin pairs (25,004 individuals) who were 33-60 years old in 1990. Attrition from the sample is due to death, migration or aging (70 years old).

## 2.2 Descriptive statistics of the twin sample

These analyses are focused on twin pairs who are both employed at year  $t$  in either the public or the private sector. Entrepreneurs and individuals with missing information are excluded. Individuals who have zero earnings (wage salaries) at year  $t$  are also excluded

---

<sup>4</sup> Identical (MZ) twins are conceived from a single fertilised egg, whereas fraternal (DZ) twins are conceived from two separate eggs. MZ twins share 100 per cent of their genes; DZ twins share 50 per cent of their genes on average.



for that particular year because these individuals are unlikely to be employed. The earnings are deflated in 2004 euros using the consumer price index. The final sample consists of 1,780 male and 2,025 female twin pairs. The data are unbalanced because not every pair is observed every year; i.e., there are gaps in the data. In total, the data include 21,116 male and 23,802 female yearly observations for the period from 1990 to 2004.

The descriptive statistics in Table 1 provides an overview of the sample by gender and sector. The share of identical twins in the sample is 35 per cent for males and 37 per cent for females. A higher share of females than males work in the public sector (44 versus 16 per cent). The average annual earnings are 6 per cent higher in the public sector for males, whereas the average pay gap for females is 1 per cent in favour of private sector employees. The individuals in the sample are 45 years old on average, but the public sector employees are slightly younger.<sup>5</sup> Consistent with the existing literature, the public sector employees are more educated and also have higher socio-economic status than private sector employees (e.g., Kanellopoulos 1997, Adamchik et al. 2000, Christofides et al. 2002). Private sector employees are, in turn, more experienced and they also report considerable longer job tenure (approximately 4 versus 12 years), which reflects the typical use of temporary contracts in public sector jobs. The classification by education field shows that public sector workers are more likely to have an educational background in teaching, health, social work and services, whereas private sector workers

---

<sup>5</sup> Earlier literature has shown that public sector employees are older than private sector employees because younger individuals typically have greater access to private sector jobs than to public sector ones, partly because they have not accumulated the experience and connections required to secure a job in the public sector (e.g., Dustmann et al. 1998, Christofides et al. 2002). The statistics drawn from my sample are thus reasonable, given the lack of young individuals in the dataset.

are more likely to have been educated in general and technology fields. Finally, a higher share of public sector employees are married or cohabiting and have children compared to their counterparts in the private sector (see also Adamchik et al. 2000, Christofides et al. 2002).

**[Insert table 1 here]**

### **3. Employment sector: genetic and environmental influences**

This section studies the impact of genetic factors and shared environment on public sector employment. First, the econometric models are outlined; then the main results and the various robustness tests are presented. The dependent variable is public sector employment by year from 1990 to 2004.

#### **3.1 Econometrical models**

Two empirical approaches are used to study to what extent public sector employment varies due to genetic and environmental factors. A classical method of addressing this question is to decompose the variation among twins and compare the outcome correlations for MZ twin pairs with those of DZ twin pairs. The first approach uses standard behavioural genetics decomposition (see, e.g., Goldberger 1979) where the genetic heritability of public sector employment is denoted as:

$$h^2 = 2(r_{MZ} - r_{DZ}) \tag{1}$$

where  $r_{MZ}$  is the correlation of public sector employment within identical twins and  $r_{DZ}$  is the corresponding correlation for fraternal twins. The fraction of variance explained by the shared environment is calculated as the difference between the correlation for the identical twins and genetic heritability:

$$c^2 = r_{MZ} - h^2 \quad (2)$$

This model relies on the following four assumptions. The first assumption is that genes and the environment have additive effects. The second is that the additive environmental influence is the same for identical and fraternal twins. The third is that assortative mating does not occur, and the fourth is that there is no correlation or interaction between genetic factors and shared environment.<sup>6</sup>

The second approach is the use of the well-known DF-model (DeFries and Fulker 1985), which relies on the same four assumptions as the standard decomposition model. The DF model produces similar estimates for shared environmental variance ( $c^2$ ) and heritability ( $h^2$ ) from the following equation that is estimated using ordinary least squares (OLS):

$$\text{Public}_{1ft} = \beta_0 + \beta_1(\text{Public}_{2ft}) + \beta_2R + \beta_3(R*\text{Public}_{2ft}) + \varepsilon_{1ft} \quad (3)$$

---

<sup>6</sup>This assumption is criticized based on the idea that individuals who inherit relatively strong genetic endowment tend to grow up in families with more favorable environments for child development (e.g., Kohler et al. 2010, Manski 2011).

This is typically referred to as the ACE specification, where A stands for additive genetics, C for a common environment and E for error. The dependent variable in the analysis is the number of years of public sector employment between 1990 and 2004. Therefore,  $\text{Public}_{1ft}$  represents the years of public sector employment for the first twin in a twin pair (family  $f$ ) at year  $t$ , and  $\text{Public}_{2ft}$  is the corresponding score variable for the second member of a twin pair. The variable for genetic relatedness is  $R$ , which yields a value of 1 for identical twins and a value of 0.5 for fraternal twins. It can be shown that  $E(\beta_1) = c^2$  and  $E(\beta_3) = h^2$ , i.e., we have estimates for heritability,  $h^2$ , and for shared environmental variance,  $c^2$ . The double-entry method employed by Cherny et al. (1992) is used here. This method involves entering each twin's score twice, once as the proband score and once as the co-twin score.

Sometimes, the correlation of the outcome variable within identical twins is more than twice the corresponding figure for fraternal twins, i.e.,  $r_{MZ} > 2r_{DZ}$ . This finding suggests that the data may be significantly influenced by additive genetic effects and that the model can yield estimates that fall within the categories  $h^2 > 1$  and/or  $c^2 < 0$ . If we assume that genetic effects need not be additive, the basic DF-model can be reformulated as follows:

$$\text{Public}_{1ft} = \beta_0 + \beta_2 R + \beta_3 (R * \text{Public}_{2ft}) + \beta_4 (D * \text{Public}_{2ft}) + \varepsilon_{1ft} \quad (4)$$

This is typically called the ADE specification, where  $D$  denotes the dominant genetic relatedness. The latter takes a value of 1 for identical twins and a value of 0.25 for fraternal twins (Waller 1994). The parameter for narrow-sense heritability is  $\beta_3$ , and the

estimated dominance effect is  $\beta_4$ . Broad-sense heritability is the sum of these two parameters,  $\beta_3 + \beta_4$ , and corresponds to the heritability estimate  $h^2$ .

If the observed variance in public sector employment depends only on the shared environment, these approaches yield estimates that are consistent with  $c^2 > 0$  and  $h^2 = 0$ . Such a finding will indicate that the propensity to enter public sector employment may increase according to a parent's occupational preferences through the roles that the child learn in the family setting, role-modelling and the transmission of job market information from the parents. In turn, if genetic factors are the more important ones, the results will be consistent with  $c^2 = 0$  and  $h^2 > 0$ . Such results will indicate that public sector employment is influenced by the genetic factors that are transmitted biologically from parents to children. It would also possible to combine these theories such that the variance in public sector employment is explained both by genetic factors and by shared environments, i.e.,  $c^2 > 0$  and  $h^2 > 0$ .

### 3.2 Interpreting the estimates for heritability: some criticism

Despite its popularity, the behavioral genetics model has been subject to criticism over the years, mainly because heritability estimates may reflect genetically induced environmental effects (see, Stenberg 2010 for a review note). In essence, the presence of gene-environment correlation<sup>7</sup>, in which the probability of an individual's being exposed to a specific environment differs based on genetics, may lead to

---

<sup>7</sup> See Plomin et al. (1977), who distinguish between three gene-environment correlation mechanisms through which genes may influence environments and indirectly generate various outcomes. These mechanisms are passive genetic influence, evocative genetic influence and active genetic influence.

misinterpretation of the *true* impact of heritability estimates. The assumption that environments are influenced by genetics, together with the possibility of gene-environment interaction, reciprocal causation and multiplier effects, suggests that estimates of  $h^2$  may be partly generated by environmental factors and may therefore often be overestimated by authors.

Most recently, the relevance of policy analysis based on variance decomposition has also been disputed (e.g., Beauchamp et al. 2011, Manski 2011). Manski (2011) argues that because  $h^2$  and  $c^2$  are only symbolic representations of hypothesised latent forces, variance decomposition does not yield estimates of political relevance or provide evidence of the relative importance of either genetic nor environmental factors. Beauchamp et al. (2011) present a more advanced view of the research on heritability and the information that it provides that may assist in policy analysis. Employing a molecular genetics framework, the researchers obtain promising results on the effects of specific alleles on the tested outcome variable (educational attainment), but they fail to replicate the results during the second stage of the study, which involves the use of data from another survey. Nevertheless, they make a cautious argument that large sample sizes must be used to identify significant genetic influences on economic variables.

### 3.3 Main results

Table 2 separately presents the correlation coefficients for identical and fraternal twins and by gender. Based on the Chow test, the correlation for years of public sector employment within identical twins is substantially higher than the corresponding figure for fraternal twins. Based on standard additive decomposition method, the estimate for

$h^2$  is 0.44 for males and 0.38 for females. The estimate for shared environments,  $c^2$ , is considerably lower at 0.01 for males and 0.11 for females. These numbers illustrate the importance of genetic factors in explaining public sector employment (at approximately 40 per cent for both genders) but also indicate that females may be influenced by factors related to shared environments.

Table 3 presents the results of the DF regression analyses, in which the standard errors are in all cases clustered by twin pairs. To save space, I only report the estimates for  $h^2$  and  $c^2$ . The estimate for heritability is large and highly significant for both genders. The estimate is 0.44 for males and 0.37 for females, and these coefficients are consistent with the correlation coefficients. Because the measure of shared environment variance is statistically zero, an alternative specification is considered in which this term is dropped from the equation (i.e., we set  $c^2 = 0$ ). This specification is titled AE and the results are reported in columns 2 and 4. Based on the Akaike information criterion, the ACE specification is preferred for females, whereas the reduced AE specification is preferred for males. The indicated measure of heritability for males is now 0.45. Under the assumptions of the DF model, the results suggest that public sector employment is heritable at a rate of approximately 40 per cent and that the effect of shared environments is statistically zero.

**[Insert tables 2-3 here]**

DeFries et al. (1985) note that their regression model can be extended to include other controls that are associated with the outcome variable. This approach is used in several econometric analyses in behavioural genetics (see also Miller et al. 1996 and 2001,

Nicolaou et al. 2008). Although they do not formally prove it, Cropanzano and James (1990) also argue that the model requires a correction for potential variables for researchers to be able to make strong statements about the effects of genetics and shared environments on the outcome variable. The additional controls considered for inclusion in this analysis are education, education field, age, socio-economic status, marital status, the presence of underage children and a dummy for having one's own house. These variables are found to be associated with public sector employment in previous studies (e.g., Ophem 1993, Kanellopoulos 1997, Adamchik et al. 2000, Christofides et al. 2002).

The results are shown in Tables 4 and 5. The Akaike information criterion indicates that the ACE specification is always preferred for females and that the reduced AE specification is preferred for males. Therefore, only the preferred model results are reported. The results indicate that once the additional control variables are held constant, the direct impact of heritability decreases from 45 to 32 per cent for males and from 37 to 29 per cent for females. Education variables – and mainly education field - play a notable part in this respect, whereas the other confounding variables do not seem to affect the heritability estimates. In other words, some of the impact of heritability on public sector employment might be influential because of its effect on educational outcomes.<sup>8</sup> That education field would have the highest impact on heritability estimates is reasonable because one's field of education substantially determines one's employment sector via

---

<sup>8</sup> See also Nicolaou et al. (2008) who study the heritability of self-employment. They reran their analysis adjusting the model for potential controls associated with self-employment, including income, education, marital status, age, ethnic group and immigrant status. They find that the extended model does little to reduce the heritability estimates for self-employment by 7 percentage points, from 0.48 to 0.41. In line with my results, Nicolaou et al. (2008) find that the decrease in heritability estimate is mainly due to inclusion of education variable into the model.



both educational and occupational interests, which are also heritable (Miller et al. 2001, Whiston et al. 2004).

**[Insert tables 4-5 here]**

### 3.4 Robustness analysis

Robustness analyses can help researchers to assess the effects of the associated regressors. Here, the DF model was rerun in two different ways. First, the continuous dependent variable (years of public sector employment) was replaced with a dummy variable that takes a value of one for the public sector and a value of zero for the private sector. Second, the analysis was restricted to the youngest cohort, those born in 1950-1957. These individuals were 33 to 40 old at the beginning of the sample period, in 1990, and were 48 to 55 years old at the end of the sample period, in 2004. Hence, this cohort represents the employees at their prime working age.

The results of the robustness tests are reported in Tables 6 and 7. For males, the estimate for shared environment is negative based on both robustness tests, which indicates that dominance effects may be present. Therefore, the ADE specification is preferred to the other specifications. The results suggest that the joint estimate (broad-sense heritability), which refers to the sum  $\beta_3 + \beta_4$ , is large and statistically significant at 39 per cent. These results are consistent with the earlier findings. The results for females, in turn, indicate the superiority of the ACE specification in both tests. As before, the effect of shared environment on the outcome variable is statistically

insignificant. The estimate for heritability is approximately 30 per cent but is statistically insignificant for the youngest cohort.

In sum, it seems that estimated regression coefficients are plausible, and the results indicate that heritability accounts for approximately 30-40 per cent of the observed variance in public sector employment. Furthermore, the analyses do give some indication of the specific mechanism that links genes to the decision to work in the public sector. Educational outcomes play an important role in this respect, accounting for approximately 10 percentage points of the direct effect of heritability on public sector employment.

**[Insert tables 6-7 here]**

## **4. Pay gaps: genetic and environmental influences**

This section adds to the literature that attempts to explain the wage differentials experienced by public and private sector employees. The special emphasis here is the role of genetics and shared environments in pay gaps. Before beginning this analysis, I briefly comment on the wage bargaining institutions for the two sectors in Finland.

### **4.1 Bargaining institutions**

Overall, Finland is heavily unionised; approximately 70 per cent of all employees belong to a trade union (see the OECD figures for recent country comparison). Among the public sector employees, the share is even higher at approximately 80 per cent.

There are three main central labour confederations with more than 70 trade unions on the employee side. On the employer side, there are four confederations, of which one is for the private sector, one is for the central government sector, one is for the local government sector and one is for the Church. In the public sector, all agreements for employees are made between the employers' confederations and the bargaining agents.

Industrial relationships are regulated by collective agreements that determine the minimum conditions for the job in question. These agreements provide the framework for branch-specific collective agreements. In all cases, the employers' associations and trade unions sign collective agreements of their own. Because collective labour contracts are binding even for non-union members, approximately 90 per cent of employees are covered by collective agreements.

Although centralised agreements have been the main mode of wage bargaining during the present decade, there has been a growing tendency towards local wage bargaining. This change has been motivated by the need to increase and ensure the competitiveness of private firms in global markets and to improve the competitive edge of the public sector in the labour market. In the public sector, the shift towards local (authority) bargaining stems from the introduction of new pay schemes that are based on job evaluations and performance appraisals. Since the beginning of 2008, the new pay system has been applied across the entire public sector, where employee remuneration includes of a job-specific and a personal pay component.

## 4.2 Econometric models

First, public-private sector pay gaps are analysed using a basic Mincerian wage equation that is augmented with a public sector dummy in the following manner:

$$\log(w_{ift}) = \beta \text{Pub}_{ift} + \gamma X_{ift} + A_{if} + \varepsilon_{ift} \quad (5)$$

where  $i$ ,  $f$  and  $t$  are indices for the individual, the family and time, respectively. Therefore,  $w_{ift}$  is the annual wage earned by individual  $i$  ( $i = 1, 2$ ) in a family  $f$  ( $f = 1, \dots, F$ ) during year  $t$  ( $t = 1990, \dots, 2004$ ).  $\text{Pub}_{ift}$  is a dummy variable that indicates whether an individual is a public or private sector worker (one for public; zero otherwise).  $X_{ift}$  is a vector of individual characteristics that includes potential work experience and its squared term, work experience in one's current job (i.e., tenure), years of education, education field (nine categories), socioeconomic status (three categories), the presence of underage children, marital status (single or non-single), house ownership, year dummies and cohort age. These independent variables are similar to those used in the standard Mincerian earnings equation. Broadly defined,  $A_{if}$  is ability and includes, for example, intelligence, access to schooling and an instructive parenting style. Finally,  $\varepsilon_{ift}$  is the iid error.

The assumption is that  $E(\text{Pub}_{ift}, X_{ift}, A_{if} | \varepsilon_{ift}) = 0$ , and ability is allowed to correlate with observed independent variables such that  $E(\text{Pub}_{ift}, X_{ift} | A_{if}) \neq 0$ . The wage equation is estimated using ordinary least squares (OLS). The parameter of interest is  $\beta$ , which measures in percentage terms how much less or more public sector workers earn (on average) than do workers in the private sector.

Because  $A_{if}$  is generally unobserved (or poorly measured), equation (5) omits this term and provides an estimate of  $\beta$  that is denoted by  $\beta_{OLS}$ . The omitted variable bias resulting from  $A_{if}$  is:

$$\beta_{OLS} = \beta + \text{Cov}(\text{Pub}_{ift}, A_{if}) / \text{Var}(\text{Pub}_{ift}) \quad (6)$$

Thus, the direction of the bias depends on the correlation between ability and public sector employment. For example, negative covariance will lead  $\beta_{OLS}$  to underestimate the true value of  $\beta$ .

The assumption is that  $A_{if}$  can be expressed as:  $A_{if} = c_f + h_{if}$  where  $c_f$  denotes the effects of a shared environment (i.e., family background effects) and  $h_{if}$  is a term for genetics (i.e., intelligence).  $c_f$  is assumed to be similar for twins raised in the same family  $f$ . However,  $h_{if}$  is assumed to be similar only for identical twins because they share 100 per cent of their genes. Because parameter  $\beta_{OLS}$  is most likely inconsistent estimator for  $\beta$ , data for twins are used here. Consider the wage equations for twins 1 and 2 raised in family  $f$ .

$$\begin{aligned} \log(w_{1ft}) &= \beta \text{Pub}_{1ft} + \gamma X_{1ft} + c_f + h_{1f} + \varepsilon_{1ft} \\ \log(w_{2ft}) &= \beta \text{Pub}_{2ft} + \gamma X_{2ft} + c_f + h_{2f} + \varepsilon_{2ft} \end{aligned} \quad (7)$$

A within-twin pair (hereafter WTP) estimator is of the form:

$$\log(w_{2ft}) - \log(w_{1ft}) = \beta(\text{Pub}_{2ft} - \text{Pub}_{1ft}) + \gamma(X_{2ft} - X_{1ft}) + (h_{2f} - h_{1f}) + (\varepsilon_{2ft} - \varepsilon_{1ft}) \quad (8)$$

where the shared environment effect  $c_f$  has been differenced out. Furthermore, if the twins are identical, then  $h_{2f} - h_{1f} = 0$ , and thus, the genetic effect can also be differenced out. The basic idea of underlying the WTP method is that return to public sector employment can be estimated controlling for the part of the unobserved heterogeneity that is due to shared environment and genetics. Equation (8) provides an estimate for  $\beta$  that is denoted by  $\beta_{WTP}$ . The assumption is that  $\beta_{WTP}$  is a consistent estimate of  $\beta$  and is less biased than  $\beta_{OLS}$ . The WTP method also conditions out the endogeneity problem of employment sector that is caused by the shared environment, genetics and the between-twins differences in vector  $X_{ift}$ .

#### 4.3 Main results

Table 8 presents the results obtained from earnings equation (5), which treats the sample of twins as a sample of individuals. The standard errors are robust to within-twin pair correlation. To determine the representativeness of the twin sample, we can compare the results with the data from Statistics Finland for the same period (from 1990 to 2004) and the same age distribution as the twin data. These check-up data are longitudinal register-based data (LRBD) for a 7 per cent random sample of the Finnish population in 2001.

The results show that the sample of twins is sufficiently representative of the sample of the general population in Finland. The findings indicate that on average, male workers in the public sector earn 7 per cent less than private sector workers. The pay gap for females is somewhat smaller at negative 4 per cent. The result for males is comparable with other Finnish evidence that reports a pay gap ranging of negative 3-5 per cent. The result for females is somewhat surprising because earlier studies find a

positive pay gap of between 1-7 per cent on average. (Korkeamäki 1999, Uusitalo 1999, Maczulskij et al. 2011).<sup>9</sup>

Table 9 contains the estimates obtained using the WTP model, equation (8). For the sake of comparison, the estimates are also reported for all twins and fraternal twins. The identical twins are identical in terms of their genetic inheritance and were raised in the same environment. The model for identical twins thus provides an estimate of the public sector wage effect that is not biased by the omission of genetics and shared environment. In contrast, the fraternal twins are not identical in their genetic inheritance but share the same environment. Thus, the estimate is biased by the omission of genetics but is not biased by the omission of shared environment.

The results for males indicate that it is important to control for shared environment and genetic variables and that when these factors are controlled for, the estimate of the return to public sector employment is zero (as compared to baseline OLS estimate of negative 7 per cent, which does not control for these factors). Interestingly, the model for fraternal twins that controls only for shared environments produces a pay

---

<sup>9</sup> Uusitalo (1999) reports a public-private sector pay gap of approximately negative 3 per cent for Finnish men using a conventional OLS method. Korkeamäki (1999) uses the Oaxaca decomposition method to estimate the fraction of the observed pay gap that is due to differences in returns to characteristics and finds that the pay gap is negative 3-5 per cent for males, whereas females receive a small pay premium. Similarly, Maczulskij et al. (2011) use the quantile regression method to estimate the pay gap at different parts of the earnings distribution and find that the pay gap for males is negative 5 per cent at the median. The corresponding pay gap for females is 7 per cent in favour of public sector employees. The typical positive pay gap for females is probably a function of the method used by Korkeamäki (1999) and Maczulskij et al. (2011), in which the working sector is treated as endogenous and selectivity is controlled in the wage equations.

gap of negative 6 per cent, which is only 1 percentage point lower than the OLS result. This finding indicates that the difference in the estimated pay gap is mainly explained by genetic factors such as intelligence.

The results for females suggest that once both the shared environment and genetic effects are held constant, the public sector wage effect is statistically significant at an 11 per cent significance level. However, the point estimate is  $-0.053$ , which indicates that the pay gap is at least economically significant at approximately 5 per cent.

**[Insert tables 8-9 here]**

#### 4.4 Robustness analysis

Two robustness checks are estimated here. In the previous estimations, individuals who had zero earnings were excluded. Therefore, it is important to determine whether the results are sensitive to these omitted observations. The first robustness test thus includes individuals who have zero earnings into the model. When working with the logarithm of income, I add one to each individual's earnings (annual earnings plus one). The results that are reported in Table 10 are comparable to the main findings. In particular, the OLS results indicate that men earn approximately 8 per cent less in the public sector, whereas female employees in the public sector receive a pay disadvantage of approximately 3 per cent. After both shared environment and genetic effects are controlled for, the pay gaps drop to zero for both sexes, except that for females, the estimate is economically significant at approximately negative 4 per cent.

In a second robustness test, the DF model is augmented with the public sector dummy and other individual controls in vector  $X_{it}$ . This approach makes it possible to



compare the returns to public sector employment between the economic model (the WTP model) and the behavioural genetics model (the DF model).<sup>10</sup> The preferred model results according to the Akaike information criterion are reported in Table 11. Columns (1) and (3) report the results obtained using the basic DF model where individual earnings are explained only by the shared environment ( $c^2$ ) and genetic factors ( $h^2$ ). Columns (2) and (4) augment the model with a public sector dummy and other individual characteristics.

The results indicate that the return of public sector employment is negative 4 per cent for females, which is comparable with the figure obtained using the WTP model. This finding provides some evidence that females receive approximately 4-5 per cent less in the public sector and that the roles of shared environment and genetics in this gap are essentially zero. The return of public sector employment for males is now negative 6 per cent, which is lower than the figure obtained using the WTP model (zero) but similar to the figure obtained using the traditional OLS model (negative 7 per cent). Thus, the results for men are not entirely robust across specifications.<sup>11</sup>

---

<sup>10</sup> See the detailed description of the model in the Appendix.

<sup>11</sup> The estimation results obtained using the DF model and the WTP may differ model for many reasons. First, the shared environment and genetic effects are explicitly estimated in the DF model and implicitly controlled for in the WTP model. Second, the shared environment control is mediated *through* the sibling's indicator variable in the DF model (see also Miller et al. 1996). Third, measurement error could be a factor. If so, the bias in the WTP estimator will be greater than that of the DF model, where there is only one variable that is prone to measurement error rather than two variables (see also Miller et al. 1996). Fourth, the WTP approach is more effective than the DF model in eliminating the potential issue of endogeneity with regard to public sector employment.

## 5. Conclusions

In the literature, public sector employment is commonly considered as an endogenous outcome variable. In recent decades, researchers have focused on the role of numerous demographic and family background factors, together with psychological features, in determining public sector participation decisions. Because many of these outcomes are mostly influenced by genetic factors, it is reasonable to consider public sector employment as an occupational choice that is actually genetically heritable. In addition, this article fills the gap in a research field that attempts to explain why public-private sector pay gaps exist. The special emphasis here is on whether the wage differentials between the sectors are attributed to 1) genetic factors such as intelligence and 2) shared environments (e.g., role-modelling provided by parents). The analyses were conducted using data on Finnish twins over a period of 15 years.

The traditional DF method was used to estimate the fraction of variation in public sector employment that is due to shared environment and genetic heritability. The findings of the analyses indicate that public sector employment is broadly influenced by genetic factors (at approximately 40 per cent), whereas the role of shared environments remains statistically zero. However, to interpret these results as a strong evidence of the superior importance of genetic factors in comparison with environmental factors would require more discussion.

According to various critics (e.g., Stenberg 2010, Manski 2011), overestimates of genetic heritability are common. All such estimates include (although to an unknown extent) different mechanisms that may hamper the interpretation of the *true* impact of heritability. Do these estimates mostly reflect the direct effects of genetic composition

on personality traits, intelligence or risk-taking behaviour? Or are they more due to indirect gene-environmental correlation effects? For example, genes might lead individuals to select environments more favourable to a certain vocational activity, which in turn will determine to a significant extent the sector in which the individual will be employed. Thus, one mechanism that may affect the relationship between genes and environment is educational (or vocational) attainment. Some studies argue that parents' genes may affect their behaviour and thereby influence the environment of the child (e.g., Plomin et al. 1977). It is therefore possible that highly educated parents who value healthy life-styles raise their children in an intelligent and health-conscious environment. Such an environment may affect the child's choice of occupation; for instance, the child might choose a career in the health industry, which is typically situated in the public sector.

Despite the critique of the behavioural genetics model, my results do provide some indication of why genes might matter so much. The findings indicate that the inclusion of education and education field within the model decreases the estimate for heritability to 30 per cent for both sexes. This finding suggests that gene-environment correlations might indeed influence the tendency of people to pursue public sector employment by affecting their educational choices. Please note that the findings of this article do not attempt to explain all of the mechanisms that links genes, the environment and public sector employment together. What this paper's DF estimations show is that public sector employment is certainly an endogenous occupational outcome. Similar statements have already been made in the literature, but the present article offers a more advanced view of this popular question.

The pay gap analysis provides results that are fundamentally informative. The OLS estimates suggest that females receive 4 per cent less in the public sector. Once both the shared environment and genetic effects are held constant, the pay gap remains at the level of negative 4-5 per cent. Thus, neither genetics nor shared environments explain the pay gap between women in the public and private sectors. The comparable WTP results for males indicate that the pay gap is zero rather than negative 7 per cent conventionally estimated. We can conclude that neglecting the control of genetics and shared environment effects in estimating the impact of public sector employment on men's earnings will yield downward biased estimates. Interestingly, the bias with regard to the pay gap is mainly caused by genetic factors. This finding could indicate that the private sector attracts relatively more able male employees than the public sector does (see also, e.g., Nickell et al. 2005) or that males who have higher risk preferences tend to select private sector employment as suggested by Hartog et al. (2002), among others.

That the present paper found that genetic factors may affect public sector participation is unsurprising given the number of occupation-related behaviors that have already been found to be heritable (Whiston et al. 2004 for a survey, Nicolaou et al. 2008). Additionally, as suggested by Manski (2011), one should not make strong policy statements based on the procedure used to decompose the observed variance in outcome variables into genetic and environmental influences. However, this article is significant within the research field intended to understand why people with different family-of-origin attributes choose public sector employment. In addition, it is important to study how these attributes affect the wage differentials between the public and private sectors; indeed, this question would have interesting implications for political science in general. The aim of pay gap analysis is to provide policy makers with information that can be

used to reduce unequal pay and develop reward strategies in public sector organisations. Such efforts should be conducted in a manner that takes into account the two main conclusions of this article. First, females receive lower pay in the public sector, and second, the private sector might attract more able male employees, making it challenging for the public sector to retain high-quality labour.

## References

- Adamchik, V., and Bedi, A. (2000) Wage Differentials in the Public and the Private Sectors: Evidence from an Economy in Transition, *Labour Economics*, **7**, 203-224.
- Beauchamp, J., Cesarini, D., Johannesson, M., van der Loos, M., Koellinger, P., Groenen, P., Fowler, J., Rosenquist, J., Thurik, A. and Christakis, N. (2011) Molecular Genetics and Economics, *Journal of Economic Perspectives*, **25**, 57-82.
- Bellante, D. and Link, A. (1981) Are Public Sector Workers More Risk Averse than Private Sector Workers?, *Industrial and Labor Relations Review*, **34**, 408-412.
- Bender, K. and Elliot, R. (2002) The Role of Job Attributes in Understanding the Public-Private Wage Differential, *Industrial Relations*, **41**, 407-421.
- Cesarini, D., Dawes, C., Johannesson, M., Lichtenstein, P. and Wallace, B. (2009) Genetic Variation in Preferences for Giving and Risk Taking, *The Quarterly Journal of Economics*, **124**, 809-842.
- Cherny, S., Cardon, L., Fulker, D. and DeFries J. (1992) Differential Heritability Across Levels of Cognitive Ability, *Behavior Genetics*, **22**, 153-162.
- Christofides, L., and Pashardes, P. (2002) Self/Paid-Employment, Public/Private Sector Selection, and Wage Differentials, *Labour Economics*, **9**, 737-762.
- Corcoran, S., Evans, W. and Schwab, R. (2004) Changing Labor Market Opportunities for Women and the Quality of Teachers, *American Economic Review*, **94**, 230-235.

- Crewson, P. (1997) Public-Service Motivation: Building Empirical Evidence of Incidence and Effect, *Journal of Public Administration Research and Theory*, **7**, 499-518.
- Cropanzano, R. and James, J. (1990) Some Methodological Considerations for the Behavioural Genetic Analysis of Work Attitudes, *Journal of Applied Psychology*, **75**, 433-439.
- DeFries, J. and Fulker, D. (1985). Multiple Regression Analysis of Twin Data, *Behavior Genetics*, **15**, 467-473.
- Dustmann, C. and van Soest, A. (1998) Public and Private Sector Wages of Male Workers in Germany, *European Economics Review*, **42**, 1417-1441.
- Goldberger, A. (1979) Heritability, *Economica*, **46**, 327-347.
- Hartog, J., Ferrer-i-Carbonell, A. and Jonker, N. (2002) Linking Measured Risk Aversion to Individual Characteristics, *KYKLOS*, **55**, 3-26.
- Johnson, W., McGue, M., Krueger, R. and Bouchard, T. (2004) Marriage and Personality: A Genetic Analysis, *Journal of Personality and Social Psychology*, **86**, 285-294.
- Kanellopoulos, C. (1997) Public-Private Wage Differentials in Greece, *Applied Economics*, **29**, 1023-1032.
- Kaprio, J., Koskenvuo, M., Artimo, M., Sarna, S. and Rantasalo, I. (1979) Baseline Characteristics of the Finnish Twin Registry: Section I: Materials, Methods, Representativeness and Results for Variables Special to Twin Studies. Kansanterveystieteen julkaisuja M47. (in Finnish)

- Karl, K. and Sutton, C. (1998) Job Values in Today's Workforce: A Comparison of Public and Private Sector Employees, *Public Personnel Management*, **27**, 515-527.
- Keller, L., Bouchard, T., Arvey, R., Segal, N. and Dawis, R. (1992) Work values: Genetic and Environmental Influences, *Journal of Applied Psychology*, **77**, 79-88.
- Kohler, H-P., Behrman, J. and Schnittker, J. (2010) Social Science Methods for Twin Data: Integrating Causality, Endowments and Heritability, University of Pennsylvania working papers.
- Korkeamäki, O. (1999) Yksityisen ja julkisen sektorin palkkoihin vaikuttavat tekijät [in Finnish], VATT-research reports No. 57, Helsinki.
- Lassibille, G. (1998). Wage Gaps Between the Public and Private Sectors in Spain, *Economics of Education Review*, **17**, 83-92.
- Lewis, G. and Frank, S. (2002) Who Wants to Work for the Government?, *Public Administration Review*, **62**, 395-404.
- Lucifora, C. and Meurs, D. (2006) The Public Sector Pay Gap in France, Great Britain and Italy, *Review of Income and Wealth*, **52**, 43-59.
- Maczulskij, T. and Pehkonen, J. (2011) Public-Private Sector Pay Gaps in Finland: A Quantile Regression Analysis, *Finnish Economic Papers*, **24**, 111-127.
- Manski, C. (2011) Genes, Eyeglasses, and Social Policy, *Journal of Economic Perspectives*, **25**, 83-94.



- Miller, P., Mulvey, C. and Martin, N. (1996) Multiple Regression Analysis of the Occupational Status of Twins: A Comparison of Economic and Behavioural Genetics Models, *Oxford Bulletin of Economics and Statistics*, **58**, 227-239.
- Miller, P., Mulvey, C. and Martin, N. (2001) Genetic and Environmental Contributions to Educational attainment in Australia, *Economics of Education Review*, **20**, 211-224.
- Nickell, S. and Quintini, G. (2002) The Consequences of the Decline in Public Sector Pay in Britain: A Little Bit of Evidence, *Economic Journal*, **112**, F107-F118.
- Nicolaou, N., Shane, S., Cherkas, L., Hunkin, J. and Spector, T.D. (2008) Is the Tendency to Engage in Entrepreneurship Genetic?, *Management Science*, **54**, 167-179.
- Ophem, H. (1993) A Modified Switching Regression Model for Earnings Differentials Between the Public and Private Sectors in Netherlands, *The Review of Economics and Statistics*, **75**, 693-709.
- Perry, J. and Wise, L. (1990) The Motivational Bases of Public Service, *Public Administration Review*, **50**, 367-373.
- Pfeifer, C. (2011) Risk Aversion and Sorting into Public Sector Employment, *German Economic Review*, **12**, 85-99.
- Plomin, R., DeFries, J. and Loehlin, J. (1977) Genotype-Environment Interaction and Correlation in the Analysis of Human Behavior, *Psychological Bulletin*, **84**, 309-322.

Plomin, R. and Spinath, F. (2004) Intelligence: Genetics, Genes, and Genomics, *Journal of Personality and Social Psychology*, **86**, 112-129.

Stenberg, A. (2010) Interpreting Estimates of Heritability - A Note on the Twin Decomposition, *Economics and Human Biology*, doi: 10.1016/j.ehb.2012.05.002.

Uusitalo, R. (1999) Return to Education in Finland, *Labour Economics*, **6**, 569-580.

Waller, N. (1994) A DeFries and Fulker Regression Model for Genetic Nonadditivity, *Behavior Genetics*, **24**, 149-153.

Whiston, S., and Keller, B. (2004) The Influences of the Family of Origin on Career Development: A Review and Analysis, *The Counseling Psychologist*, **32**, 493-568.

**Table 1.** Summary statistics by gender and sector

	Men		Women	
	Public	Private	Public	Private
log(annual earnings)	10.32	10.26	9.90	9.91
Age, years	44.2	44.8	44.0	45.0
Education, years	14.2	11.8	12.4	11.1
Work experience, years	17.9	23.1	19.1	24.1
Tenure, years	4.4	12.3	4.3	11.9
Socio-economic status, dummy				
Upper-level employee	0.45	0.18	0.17	0.08
Lower-level employee	0.32	0.22	0.61	0.51
Manual worker	0.23	0.60	0.22	0.41
Education field, dummy				
General or unknown	0.21	0.42	0.28	0.52
Education	0.04	0.00	0.03	0.00
Humanistic and arts	0.06	0.00	0.03	0.01
Business and social sciences	0.17	0.08	0.18	0.21
Natural sciences	0.04	0.02	0.01	0.01
Technology	0.17	0.43	0.03	0.11
Forestry and agriculture	0.03	0.02	0.01	0.01
Health and social work	0.11	0.01	0.32	0.03
Services	0.17	0.02	0.11	0.10
Non-single, dummy	0.84	0.79	0.71	0.70
Children, dummy	0.65	0.55	0.58	0.49
Own house, dummy	0.77	0.83	0.81	0.80
Number of observations	3,387	17,729	10,540	13,262
Total number of observations	21,116		23,802	
MZ twins, dummy	0.35		0.37	

**Table 2.** Correlation coefficients

	$r_{MZ}$	$r_{DZ}$	$h^2$	$c^2$	Chow-test: Chi2-value	p-value
Men	0.45	0.23	0.44	0.01	70.03	$p < 0.01$
Women	0.49	0.30	0.38	0.11	123,92	$p < 0.01$

**Table 3.** ACE and AE model results

	Men		Women	
	ACE (1)	AE (2)	ACE (3)	AE (4)
$c^2$	.01 (.12)		.12 (.08)	
$h^2$	.44 (.17)***	.45 (.06)***	.37 (.11)***	.52 (.04)***
$R^2$	.10	.10	.15	.15
AIC	95,159.59	95,157.71	123,156.00	123,196.80
Number of obs.	21,116		23,802	

Notes: \*\*\* denotes statistical significance at least at the 1% level. Model with smallest AIC value is preferred. Standard errors are in parentheses.

**Table 4.** AE model augmented with individual controls, men

	(1)	(2)	(3)	(4)	(5)	(6)
$h^2$	.45 (.06)***	.41 (.06)***	.33 (.05)***	.32 (.05)***	.32 (.05)***	.32 (.05)***
Education years		√	√	√	√	√
Education field			√	√	√	√
Age				√	√	√
Socioeconomic status					√	√
Demographic factors						√
$R^2$	.10	.15	.27	.28	.28	.29
Number of obs.	21,116	21,116	21,116	21,116	21,116	21,116

Notes: \*\*\* denotes statistical significance at least at the 1% level. Standard errors are in parenthesis.

**Table 5.** ACE model augmented with individual controls, women

	(1)	(2)	(3)	(4)	(5)	(6)
$c^2$	.12 (.08)	.09 (.08)	.10 (.07)	.07 (.07)	.07 (.07)	.06 (.07)
$h^2$	.37 (.11)***	.37 (.11)***	.30 (.10)***	.29 (.10)***	.29 (.10)***	.29 (.10)***
Education		√	√	√	√	√
Education field			√	√	√	√
Age				√	√	√
Socioeconomic status					√	√
Demographic factors						√
$R^2$	.10	.17	.25	.28	.28	.28
Number of obs.	23,802	23,802	23,802	23,802	23,802	23,802

Notes: \*\*\* denotes statistical significance at least at the 1% level. Standard errors are in parenthesis.

**Table 6.** ACE, ADE and AE model results: dependent variable is public sector dummy

	Men			Women	
	ACE	AE	ADE	ACE	AE
	(1)	(2)	(3)	(4)	(5)
$c^2$	-.08 (.10)			.04 (.07)	
$h^2$	.47 (.14)***	.37 (.05)***	.39 (.06)***	.31 (.10)***	.37 (.03)***
$R^2$	.07	.07	.07	.07	.07
AIC	16,086.05	16,099.11	16,086.05	32,484.24	32,487.39
Number of obs.	21,116			23,802	

Notes: \*\*\* denotes statistical significance at least at the 1% level. Model with smallest AIC value is preferred.

Standard errors are in parenthesis.

**Table 7.** ACE, ADE and AE model results: individuals born in 1950-1957

	Men			Women	
	ACE	AE	ADE	ACE	AE
	(1)	(2)	(3)	(4)	(5)
$c^2$	-.23 (.18)			.18 (.15)	
$h^2$	.61 (.26)***	.33 (.09)***	.39 (.11)***	.26 (.20)	.49 (.07)***
$R^2$	.07	.07	.07	.07	.07
AIC	31,079.46	31,116.90	31,079.46	36,849.59	36,875.52
Number of obs.	6,728			7,072	

Notes: \*\*\* denotes statistical significance at least at the 1% level. Model with smallest AIC value is preferred.

Standard errors are in parenthesis.

**Table 8.** Results for LRBD data and twin data: all subjects treated as individuals

	Men		Women	
	LRBD	Twin data	LRBD	Twin data
	(1)	(2)	(3)	(4)
Public sector	-.082 (.005)***	-.072 (.021)***	-.025 (.004)***	-.037 (.013)***
Education years	.081 (.001)***	.060 (.005)***	.071 (.001)***	.065 (.006)***
Work experience	.021 (.001)***	.014 (.002)***	.017 (.001)***	.014 (.002)***
Work exper_squared	-.036 (.002)***	-.030 (.006)***	-.036 (.001)***	-.030 (.005)***
Tenure	.012 (.000)***	.010 (.001)***	.011 (.000)***	.009 (.001)***
Socio-economic status				
Lower-level employee	-.231 (.007)***	-.245 (.025)***	-.144 (.004)***	-.157 (.026)***
Manual worker	-.248 (.005)***	-.294 (.028)***	-.277 (.005)***	-.240 (.028)***
Education field				
Teaching	-.306 (.035)***	-.192 (.079)**	-.235 (.014)***	-.174 (.067)***
Humanistic and arts	-.338 (.022)***	-.346 (.070)***	-.284 (.015)***	-.350 (.068)***
Business and social sciences	-.157 (.010)***	-.063 (.041)	-.197 (.007)***	-.155 (.030)***
Natural sciences	-.198 (.018)***	-.131 (.064)**	-.136 (.020)***	-.117 (.067)*
Technology	-.194 (.006)***	-.102 (.026)***	-.198 (.007)***	-.164 (.033)***
Forestry and agriculture	-.275 (.012)***	-.224 (.038)***	-.202 (.018)***	-.211 (.060)***
Health and social work	-.006 (.018)	.132 (.059)**	-.047 (.006)***	.005 (.028)
Services	-.051 (.010)***	.046 (.037)	-.163 (.006)***	-.158 (.027)***
Children	.024 (.004)***	.027 (.013)**	-.047 (.003)***	-.060 (.011)***
Non-single	.078 (.005)***	.061 (.015)***	-.040 (.003)***	-.065 (.011)***
Own house	.065 (.005)***	.043 (.016)***	.036 (.004)***	.031 (.011)***
Cohort + year dummies	yes	yes	yes	yes
R <sup>2</sup>	.31	.36	.24	.23
Number of obs.	304,403	21,116	365,890	23,832

Notes: \*\*\*, \*\*, \* denotes statistical significance at least at the 1, 5 and 10 % significance level. Standard errors are in parenthesis and are robust to within-individual correlation for LRBD and within twin-pair correlation for twin data.

**Table 9.** Results for twin data: within-twin pairs

	Men			Women		
	All twins (1)	DZ twins (2)	MZ twins (3)	All twins (4)	DZ twins (5)	MZ twins (6)
Public sector	-.042 (.028)	-.061(.034)*	.009 (.046)	-.029 (.018)	-.019 (.022)	-.053 (.033)
Education years	.052 (.008)***	.053 (.009)***	.051 (.019)***	.044 (.008)***	.056 (.010)***	.015 (.014)
Work experience	.011 (.004)**	.010 (.005)**	.014 (.010)	.011 (.003)***	.010 (.003)***	.011 (.004)**
Work exper_squared	-.023 (.013)*	-.019 (.014)	-.037 (.030)	-.026 (.008)***	-.023 (.010)**	-.033 (.013)**
Tenure	.008 (.001)***	.008 (.001)***	.008 (.002)***	.009 (.001)***	.009 (.001)***	.010 (.001)***
Socio-economic status						
Lower-level employee	-.165 (.030)***	-.175 (.038)***	-.141 (.045)***	-.136 (.036)***	-.125 (.043)***	-.144 (.062)**
Manual worker	-.214 (.033)***	-.229 (.040)***	-.167 (.053)***	-.239 (.038)***	-.237 (.046)***	-.212 (.067)***
Education field						
Teaching	-.104 (.079)	-.058 (.094)	-.212 (.131)	-.040 (.076)	-.084 (.089)	.072 (.133)
Humanistic and arts	-.350 (.107)***	-.394 (.134)***	-.275 (.177)	-.177 (.071)**	-.208 (.092)**	-.117 (.108)
Business and social sciences	-.066 (.047)	-.078 (.056)	-.022 (.090)	-.092 (.042)**	-.100 (.056)*	-.068 (.061)
Natural sciences	-.076 (.096)	-.018 (.119)	-.204(.167)	.151 (.092)	.095 (.099)	.302 (.239)
Technology	-.093 (.037)**	-.081 (.043)*	-.141 (.078)*	-.122 (.041)***	-.149 (.052)***	-.035 (.066)
Forestry and agriculture	-.155 (.057)***	-.145 (.080)**	-.181 (.100)*	-.133 (.074)*	-.065 (.098)	-.284 (.085)***
Health and social work	.069 (.077)	.114 (.081)	-.033 (.154)	.055 (.040)	.036 (.050)	.116 (.069)*
Services	-.001 (.048)	.043 (.056)	-.083 (.090)	-.113 (.039)***	-.130 (.050)***	-.066 (.065)
Children	.038 (.016)**	.035 (.019)*	.046 (.028)	-.082 (.017)***	-.082 (.023)***	-.077 (.024)***
Non-single	.043 (.019)**	.057 (.024)**	.014 (.035)	-.069 (.015)***	-.072 (.019)***	-.053 (.027)**
Own house	.034 (.017)*	.043 (.022)**	.015 (.029)	.028 (.018)	.034 (.021)	.015 (.031)
Cohort + year dummies	yes	yes	yes	yes	yes	yes
R <sup>2</sup>	.14	.17	.09	.10	.12	.08
Number of obs.	10,558	6,893	3,665	11,901	7,507	4,394

Notes:\*\*\*, \*\*, \* denote statistical significance at least at the 1%, 5% and 10% significance level. Standard errors are in parenthesis and are robust to within-twin pair correlation



**Table 10.** Results for twin data, robustness tests

	Men			Women		
	All twins treated as individuals (1)	Within DZ twins (2)	WithinMZ twins (3)	All twins treated as individuals (4)	Within DZ twins (5)	Within MZ twins (6)
Public sector	-.086 (.030)***	-.070(.046)	.023 (.058)	-.034 (.016)**	-.020 (.030)	-.044 (.036)
Education years	.061 (.006)***	.058 (.011)***	.075 (.025)***	.069 (.007)***	.059 (.012)***	.015 (.015)
Work experience	.013 (.003)***	.009 (.006)	.015 (.013)	.014 (.002)***	.013 (.004)***	.016 (.006)***
Work exper_squared	-.026 (.008)***	-.017 (.016)	-.044 (.035)	-.029 (.006)***	-.028 (.013)**	-.044 (.016)***
Tenure	.010 (.001)***	.008 (.002)***	.007 (.002)***	.009 (.001)***	.009 (.001)***	.009 (.002)***
Socio-economic group						
Lower-level employee	-.230 (.030)***	-.172 (.043)***	-.062 (.070)	-.159 (.030)***	-.129 (.055)**	-.153 (.064)**
Manual worker	-.286 (.033)***	-.212 (.048)***	-.108 (.075)	-.248 (.033)***	-.237 (.059)***	-.229 (.071)***
Education field						
Teaching	-.216 (.114)*	-.173 (.164)	-.297 (.154)*	-.160 (.069)**	.013 (.126)	.039 (.138)
Humanistic and arts	-.329 (.089)***	-.406 (.194)**	-.039 (.335)	-.331 (.071)***	-.073 (.121)	-.095 (.131)
Business and social sciences	-.052 (.047)	-.080 (.067)	-.108 (.107)	-.158 (.038)***	-.088 (.074)	-.088 (.067)
Natural sciences	-.171 (.105)	-.181 (.194)	-.346(.197)*	.137 (.080)*	.040 (.139)	.274 (.240)
Technology	-.105 (.032)***	-.118 (.057)**	-.254 (.111)**	-.178 (.039)***	-.157 (.065)**	-.049 (.069)
Forestry and agriculture	-.226 (.048)***	-.184 (.077)**	-.264 (.111)**	-.168 (.063)***	.041 (.155)	-.279 (.091)***
Health and social work	.178 (.065)***	.131 (.098)	-.145 (.188)	.003 (.035)	.040 (.067)	.062 (.081)
Services	.051 (.048)	.001 (.070)	-.189 (.108)*	-.149 (.032)***	-.121 (.068)*	-.081 (.071)
Children	.014 (.017)	.015 (.027)	.033 (.033)	-.064 (.014)***	-.057 (.027)**	-.056 (.027)**
Non-single	.111 (.024)***	.144 (.040)***	.059 (.042)	-.051 (.014)***	-.101 (.023)***	-.071 (.032)**
Own house	.056 (.023)**	.057 (.035)*	.019 (.040)	.027 (.015)*	.028 (.030)	.011 (.035)
Cohort + year dummies	yes			yes		
R <sup>2</sup>	.15	.06	.04	.13	.06	.05
Number of obs.	21,266	6,948	3,685	23,920	7,551	4,409

Notes:\*\*\*, \*\*, \* denote statistical significance at least at the 1%, 5% and 10% significance level. Standard errors are in parenthesis and are are robust to within-twin pair correlation

**Table 11.** DF model augmented with individual controls

	Men		Women	
	ACE	ADE	AE	ADE
	(1)	(2)	(3)	(4)
$c^2$	.153 (.078)*	-	-	-
$h^2$	.327 (.115)***	.239 (.039)***	.300 (.029)***	.136 (.032)***
Public dummy		-.062 (.020)***		-.037 (.013)***
Education years		.052 (.005)***		.064 (.005)***
Work experience		.012 (.002)***		.012 (.002)***
Tenure		.010 (.001)***		.009 (.001)***
Other controls <sup>a</sup>		√		√
$R^2$	.15	.38	.05	.29
Number of obs.	21,116	21,116	23,802	23,802

Notes: \*\*\*, \*\*, \* denote statistical significance at least at the 1%, 5% and 10% level. Standard errors are in parenthesis.

<sup>a</sup>Other controls include: cohort, field of education (9 categories), presence of children under 18 year old, marital status, socioeconomic status (3 categories), dummy for own house, year and cohort dummies

## Appendix

### A1. Extended DF-regression analysis

In their study using data for Australian twins, Miller et al. (1996) compared the returns to education between the economic model (the WTP model) and the behavioural genetics model (the DF model). The latest literature also attempts to integrate the economic and behavioural genetics approaches in the analysis of twins (e.g., Kohler et al. 2010). Following Miller et al. (1996), I work with a DF model that is augmented with a public sector dummy and other individual background characteristics. The resulting equation is as follows:

$$\ln(w_{1ft}) = \beta_0 + \beta_1 \ln(w_{2ft}) + \beta_2 R + \beta_3 (R * \ln(w_{2ft})) + \beta_4 \text{Pub}_{1ft} + \beta_5 X_{1ft} + \varepsilon_{1ft} \quad (1)$$

where  $w_{1ft}$  is the annual wage earned by the first twin in a twin pair (family  $f$ ) in year  $t$  and where  $w_{2ft}$  is annual wage earned by a second member of the twin pair in year  $t$ .  $R$  is genetic relatedness, which takes a value of 1 for identical twins and a value of 0.5 for fraternal twins.  $\text{Pub}_{1ft}$  is an indicator of whether an individual is a public or a private sector worker, and  $X_{1ft}$  is a vector of other individual characteristics. The double-entry method is used here; the twins are entered as two observations. The parameter  $\beta_1$  denotes the impact of a shared environment on wages ( $c^2$ ), and parameter  $\beta_3$  denotes the impact of genetics on wages ( $h^2$ ). The parameter of interest is  $\beta_4$ , which reflects the public sector wage effect, i.e., the wage gap between public and private sector employees.