

Using Firm-Level Data to Assess Gender Wage Discrimination in the Belgian Labour Market

D. Borowczyk Martins and V. Vandenberghe

Discussion Paper 2010-7

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain



UCL

Using Firm-Level Data to Assess Gender Wage Discrimination in the Belgian Labour Market^{*}

D. Borowczyk Martins[§] and V. Vandenberghe[£]

Abstract

In this paper we explore a matched employer-employee data set to investigate the presence of gender wage discrimination in the Belgian private economy labour market. We identify and measure gender wage discrimination from firm-level data using a labour index decomposition pioneered by Hellerstein and Neumark (1995), which allows us to compare direct estimates of a gender productivity differential with those of a gender labour costs differential. We take advantage of the panel structure of the data set and identify gender wage discrimination from within-firm variation. Moreover, inspired by recent developments in the production function estimation literature, we address the problem of endogeneity in input choice using a structural production function estimator (Levinsohn and Petrin, 2003). Our results suggest that there is no gender wage discrimination inside private firms located in Belgium.

JEL Classification: J24, C52, D24

Keywords: labour productivity; wages; gender discrimination; structural production function estimation; panel data.

^{*} Funding for this research was provided by the Belgian Federal Government - SPP Politique scientifique, programme "Société et Avenir", *The Consequences of an Ageing Workforce on the Productivity of the Belgian Economy*, research contract TA/10/031A.

[§] Department of Economics, University of Bristol.

[£] Corresponding author. Economics Department, IRES, Economics School of Louvain (ESL), Université catholique de Louvain (UCL), 3 place Montesquieu, B-1348 Belgium email : vincent.vandenberghe@uclouvain.be. We would like to thank Hylke Vandenbussche for her helpful comments and suggestions on a previous version of this paper.

1. Introduction

Evidence of substantial average earning differences between men and women— what is often termed the gender pay gap — is a systematic and persistent social outcome in the labour markets of most developed economies. This social outcome is often perceived as inequitable by a large section of the population and it is generally agreed that its causes are complex, difficult to disentangle and controversial (Cain, 1986). In 1999, the gross pay gap between women and men in the EU-27 was, on average, 16% (European Commission, 2007), while in the U.S. this figure amounted to 23.5% (Blau and Kahn, 2000). Belgian statistics (Institut pour l'égalité des Femmes et des Hommes, 2006) suggest gross monthly gender wage gaps ranging from 30% for white-collar workers to 21% for blue-collar workers.¹

Although historically decreasing the gender pay gap, and particularly the objective of further reducing its magnitude, remains a central political objective in governments' agendas both in Europe and in the U.S.² The gender pay gap provides a measure of what Cain (1986) considers the practical definition of gender discrimination. In Cain's conceptual framework gender discrimination, as measured by the gender pay gap, is an observed and quantified outcome that concerns individual members of a minority group, women, and that manifests itself by a lower pay with respect to the majority group, men.

From an economic point of view, gender wage discrimination implies that equal labour services provided by equally productive workers have a sustained price/wage difference.³ This question has motivated the emergence of diverse concepts and theories of wage discrimination. Starting with Becker (1957) several theoretical models have been proposed to describe the emergence and persistence of wage discrimination under diverse economic settings. The development of a theoretical literature on gender wage discrimination was accompanied by empirical work devoted to testing the theoretical predictions of the models and to the measurement of some concept of gender wage discrimination. We briefly describe the most important theories of gender discrimination in the labour market and the main empirical approaches to the measurement of gender wage

¹ These are figures for the private sector. The gap in the public sector is only 5%.

² See European Commission (2007) for an assessment of the gender pay gap in the European Union member states and Blau and Kahn (2000), for a comprehensive analysis of the evolution of the gender pay gap in the U.S.

³ In this paper, we will refer to labour costs differences and assume that they are good proxies for wages/earnings.

discrimination in Section 2.

In this paper we measure, and test for, the presence of gender wage discrimination (as traditionally defined by economists) in the Belgian labour market by employing a methodological approach, pioneered by Hellerstein & Neumark (1995), using a large data set that matches firm-level data, retrieved from Belfirst⁴, with data from Belgian's Social Security register containing detailed information about the characteristics of the employees in those firms. This methodological approach uses firm-level data to identify and measure gender wage discrimination as the *gap* between a measure of women's compensation relative to men's (the gender wage differential)⁵ and a measure of women's productivity relative to men's (the gender productivity differential).⁶

Its main advantages over competing methodologies (see Section 2) are essentially two. First, it *provides a direct measure of gender productivity differences* that can be subsequently compared to a measure of gender labour costs differences, thereby identifying gender wage discrimination. Second, it measures, and tests for the presence of, a concept of *market-wide* gender wage discrimination. Hellerstein & Neumark's methodology has also been used to test other wage formation theories, most notably those investigating the relationship between wages and productivity along age profiles, *e.g.* Hellerstein & Neumark (1995). Extensions of the basic methodology include enlarging the scope of workers characteristics, such as age, race and marital status, *e.g.* Hellerstein *et al.* (1999) or Vandenberghe & Waltenberg (2010), and the consideration of richer data sets regarding employee information, *e.g.* Crépon, Deniau & Pérez-Duarte (2002). In this paper, we will focus on gender and also the interaction between gender and the worker's blue-vs. white-collar status.⁷

From the econometric standpoint, recent developments of Hellerstein & Neumark's methodology have tried to improve the estimation of the production function by the adoption of alternative

⁴ <http://www.bvdep.com/en/bel-first.html>

⁵ Our measure exploits labour costs data (that include gross wage and social security contributions) which are very good proxy of what employees get paid.

⁶ As to the terminology used in the paper, the reader should bear in mind that the term "differential" designates the productivity (or labour costs) differences between women and the reference (i.e. men); whereas the term "gap" refers to the difference between the productivity and the labour costs differentials characterizing women vis-à-vis men.

⁷ Historically in Belgium, white collars (or "employees") were those performing work that requires predominantly mental rather than physical effort (presumably educated people thus), whereas the blue collars (or "workmen") were employed in manual/ unskilled labour. But that distinction has partially lost its relevance, particularly for the white-collar group that now encompasses a rather heterogeneous set of activities and levels of education). The distinction also largely recoups separate industrial relation arrangements (different rights and obligations in terms of notice period, access to unemployment insurance benefits...).

strategies to deal with potential heterogeneity bias (unobserved time-invariant determinants of firms' productivity) and simultaneity bias (endogeneity in input choice in the short run that include the gender mix of the firm). Aubert & Crépon (2004) control for the heterogeneity bias using a «within» transformation, thereby identifying gender wage discrimination from within-firm variation, and deal with the simultaneity bias by estimating Arellano & Bond's (1991) GMM (Generalized Method of Moments) estimator. Dostie (2006) alternatively controls for the endogeneity in input choice by applying Levinsohn and Petrin's (2003) structural production function estimator and takes into account both firm and workplace heterogeneity in the model of wage determination.

We follow the most recent applications of Hellerstein & Neumark's methodology and explore within-firm variation provided by panel data to identify gender wage discrimination. Next, we deal with potential endogeneity in input choice by implementing Levinsohn and Petrin's (henceforth LP) (2003) intermediate good proxy approach that we implement using information on firms' varying level of intermediate consumption.⁸

Finally is important to stress that we possess (and make systematic use of) firm-level information on the *total number of hours worked annually*. We divide the latter by the number of employees (full-time or part-time ones indistinctively) and use the result (average hours worked) as a control variable for both the production and the labour cost equations. There is evidence in our data that average hours worked is negatively correlated with the share of female work: something that reflects women's higher propensity to work part-time, but that crucially needs to be controlled for to properly capture the productivity (and labour costs) effect of changes in the share of female work.

Our preferred estimates indicate that the cost of employing women⁹ is 6 percentage points lower than that of men, pointing at a wage differential of similar magnitude. But on average, women's collective contribution to a firm's value added (or productivity) is estimated to be about 6 to 12 percentage points lower than that the group of male workers. The key result of the paper, however, is that *we cannot not reject the hypothesis that the estimated gender labour costs/wage differential is equal to the estimated gender productivity differential*. Our implementation of a Wald test of equality does not lead us to reject the null hypothesis of equality between these two differentials.

⁸ It is calculated here as the differences between the firm's turnover (in nominal terms) and its net value-added. It reflects the value of goods and services consumed or used up as inputs in production by enterprises, including raw materials, services bought on the market.

⁹ And presumably their wage.

The tentative conclusion is that, for private for-profit firms based in Belgium, productivity differences between male and female workers fully account for labour costs differences.

Our labour cost estimates are consistent with evidence obtained in previous studies of the gender pay gap in the Belgian labour market (Meulders & Sissoko, 2002), in the sense that they systematically point at lower pay for women. But our work adds new results to previous evidence for two reasons mainly. First, because we use firm-level data we are also able to estimate gender productivity differences and show that firm employing more women tend to generate less value added *ceteris paribus*. Second, by estimating labour costs and productivity equations simultaneously we are able to show that there is no statistically significant *gap* between the gender labour cost differential and the gender productivity differential: something that we interpret at the absence of wage discrimination.

The rest of the paper is organised in the following way. In Section 2 we briefly describe the most important theories of gender discrimination and review alternative empirical approaches to Hellerstein & Neumark's methodology. Section 3 describes the methodological approach: the labour-quality-index-augmented production function and labour costs equation specifications are presented in subsection 3.1; subsection 3.2 provides a description of the econometric model that underlies our empirical analysis; finally, the model of firms' behaviour underlying LP's production function estimator is sketched in subsection 3.3. Section 4 describes the data and presents summary statistics. In Section 5 we present, discuss and interpret the results of our preferred econometric specifications. Section 6 summarizes and concludes our analysis.

2 Literature

This section briefly describes the most important theories of gender discrimination related to the labour market and the empirical approaches that have been used to quantify gender wage discrimination.

2.1 Theories and Concepts of Economic Discrimination.

In framing the theoretical discussion on economic discrimination it is convenient to distinguish *i*) concepts of economic discrimination (the way is defined) from *ii*) theories of economic

discrimination (the mechanisms that cause wage discrimination or that are likely to counteract this phenomenon).

We start with the concepts, namely gender wage discrimination and gender employment discrimination. Gender wage discrimination concerns *the observation* of sustained differences in pay between men and women with equally productive capacity. Some of its constituents deserve attention. First, its focus is individual differences in pay of members of different groups for the remuneration of some service provided in a formal labour market. Second, the content of the term "equal productive capacity" requires substantiation: it refers to the output of a broad definition of some material or physical production process, which therefore excludes potential psychic disutility to employers, workers or costumers associated with the provision of those services. Gender employment discrimination concerns a differential treatment of women with respect to men in hiring and promotion decisions by employers.

We now turn to economic theories of discrimination, focusing on their prediction regarding the prevalence and persistence of wage discrimination. The neoclassical literature identifies three mechanisms that generate wage differences above productivity differences between women and men in the labour market.

The first and most famous theory of economic discrimination is due to Becker (1957). In Becker's model, employers hold a 'taste for discrimination,' meaning that there is a disutility to employing minority workers (*e.g.* women). Hence, minority workers may have to 'compensate' employers by being more productive at a given wage or, equivalently, by accepting a lower wage for identical productivity. However, the central prediction derived from Becker's various models is that the efficiency costs associated with prejudiced preferences by employers would eliminate wage discrimination in the long run.¹⁰

However, taste-based discrimination theories lead to substantially different predictions when *search friction* environments are analyzed. The central intuition is that under imperfect information about jobs, employees, employers and costumers, the segregation and free-entry mechanisms (in the case of employer discrimination) that drive out economic discrimination in Becker's model may be substantially impaired, so that wage discrimination will likely survive. In a setting with prejudiced

¹⁰ As Heckman (1998) points out, this corresponds to the common misinterpretation of Becker's model. Indeed, for market discrimination to disappear in the long run, either the number of non-discriminatory employers is sufficiently large to absorb all the minority group workers, or the supply of entrepreneurs is perfectly elastic in the long run at zero price.

costumers, Borjas & Bronars' (1989) conclude that wage discrimination for low-skilled self-employed workers of the minority group relative to the majority group is sustainable in the long run. Similarly, Sasaki (1999) shows that wage discrimination is sustainable in the long run when co-workers rather than employers discriminate against the minority group. Finally, Bowlus & Eckstein (2002) and Rosén (2003) show, under diverse assumptions, that when employers are prejudiced wage discrimination may not be eliminated in the long run.

A second discrimination mechanism is identified by theories of statistical discrimination, first presented by Arrow (1972) and Phelps (1972). These theories describe how imperfect *information about workers'* productivity and turnover propensity may generate group discrimination in a competitive setting where discriminating by membership to some group provides a cheap screen to employers. A first class of models stress the role of prior beliefs about group productivity and turnover propensity differences, leading to biased hiring and pay decisions. Work by Coate and Loury (1993b) has shown that statistical discrimination can lead to an equilibrium where an otherwise equally skilled minority group ends up with different levels of skills due to employers' prior beliefs about group skills differences. A second set of models (*e.g.* Aigner and Cain, 1977) highlights statistical discrimination that is generated by differential reliability of the signal supplied by each group. In the latter case this «formulation may be viewed as redefining the productivity of workers to include both the workers' physical productivity and the information workers convey about it» (Cain, 1986). Statistical discrimination theories are thus generically consistent with an outcome of wage discrimination, but, as information about the productivity of the individual employer is revealed, non-discriminatory employers should adjust wages to productivity, thereby eliminating wage discrimination. In this respect, the theoretical prediction is somewhat similar to that of Becker's taste-based discrimination theories.

A third discriminatory mechanism in the labour market is known as *the crowding hypothesis*, and was first formalized in Bergmann (1971). Suppose that, for some reason — be it collective discriminatory action or individual employer taste-based discrimination (*e.g.* Bergmann, 1974) — the minority group employment opportunities are restricted to a specific set of occupations. Then, if the size of the minority group is large enough relative to the employment opportunities in the set of specific occupations, two effects would come about. First, labour market clearance for the specific occupation would entail a reduction in productivity, and thus wages, of the employed minority group. Second, under the assumption of equally productive capacity of the two groups, the opportunity cost of the minority group would be lower with respect to the majority group. While

the first effect does not entail wage discrimination but only lower productivity and wages for the minority group, the second effect can generate wage discrimination in the non-segregated occupations.

Beyond theories of gender wage and employment discrimination, and consequently beyond the focus of this paper, research efforts have also been directed at investigating the impact of group differences in preferences and skills in labour market outcomes. These models rationalize observed differences in pay by hypothesizing differences between the minority and majority groups with respect to preferences for market versus non-market work, leisure or occupations, differences in comparative advantage and differences in human capital investment (Altonji & Blank, 1999).

2.2 Empirics of Gender Economic Discrimination

The focus of most of the empirical literature on gender wage discrimination has been on identifying and measuring gender discrimination rather than testing the theoretical predictions of some specific theory of discrimination. The standard empirical approach to the measurement of gender wage discrimination consists of estimating wage equations and applying Oaxaca (1973) and Blinder (1973) *decomposition methods*. In wage equations, wage discrimination is measured as the average mark-up, on some measure of individual compensation, associated to the membership to the minority group, controlling for individual productivity-related characteristics. In Blinder-Oaxaca decomposition method the difference in the average wage of the minority group relative to the majority group is explained by what Beblo *et al.* (2003) call the endowment effect (*i.e.* the effect of differing human capital endowments, diploma, experience but also ability) and the remuneration effect (*i.e.* different remunerations of the same endowments). And the remuneration effect has been traditionally interpreted as a measure of wage discrimination in the labour market.

The main shortcoming of this approach is that its identification strategy relies on the assumption that individuals are homogeneous in any productivity-related characteristic that is not included in the set of variables describing individuals' endowment. Two problems, one theoretical and another empirical, emerge. First, the researcher has to choose a set of potential individual productivity-related characteristics (diploma, experience, ability...). Second, he needs to find or create appropriate measures of those characteristics. While the second problem is becoming more manageable with the recent availability of rich individual-level data sets, the first problem can never be fully solved without using some measure of individual productivity. Furthermore, insofar as discrimination affects individual choices regarding human capital decisions or occupational

choices, the measure of discrimination obtained from wage equations will likely understate discrimination (Altonji & Blank, 1999).

Studies of narrowly-defined occupations and audit studies attempt to provide escape routes from these problems. Studies of narrowly-defined occupations estimate male and female wage differentials in specific occupations assuming that sector-specificity is sufficient to eliminate the heterogeneity in workers productivity-related characteristics (Gunderson, 2006). In some cases direct measures of productivity are used to compare estimates of wage and productivity differentials. In our view, this approach suffers from two drawbacks. First, assuming away the omitted-variable bias is never fully satisfactory from the methodological point of view. Second, the identification of gender discrimination is subject to sector- and occupation-specific biases, *e.g.* presence of rents that allow employers to indulge in gender discrimination etc. Audit studies, *e.g.* Neumark (1996), directly test for employment rather than wage discrimination by comparing the probability of being interviewed and the probability of being hired of essentially identical individuals aside from the membership to the minority group. Audit studies also face serious empirical challenges in ensuring that their methodological requirements are satisfied (*e.g.* guaranteeing a large number of testers, auditors homogeneity etc.). More importantly, audit studies do not identify employment discrimination occurring at the market level, indeed Heckman (1998) notes that «a well-designed audit study could uncover many individual firms that discriminate, while at the same time the marginal effect of discrimination on the wages of the employed workers could be zero».

As we mentioned in the introductory section, in this paper we implement an empirical methodology that involves obtaining estimates of firm-level *direct measures of gender productivity and wage differentials* via, respectively, the estimation of a production function and a labour costs equation both expanded by the specification of a labour-quality index. Under proper assumptions (see Section 3.1) the comparison of these two estimates provides a direct test for gender wage discrimination. One advantage of this setting is that it does not rely on productivity indicators taken at the individual level, which are known to be difficult to measure with precision, but rather at the aggregate level, namely, for groups of workers.

Moreover, because this approach uses information about firms of all sectors of the economy it properly measures, and tests for, a concept of market-wide gender discrimination. Therefore, Hellerstein & Neumark's methodology addresses some of the main identification problems of the existing empirical methodologies. Of course, in spite of its power Hellerstein & Neumark's gender

discrimination test is not bullet-proof. However, compared to Oaxaca-Blinder decomposition based on wage equations, it does not identify as gender discrimination gender wage differences that are explained by gender productivity differences.

3 Econometric modelling and methodology

3.1 Specification of the Productivity and Wage Differentials

In order to estimate gender-productivity (and similarly gender-wage profiles), following many authors in this area, we first consider a Cobb-Douglas production function (Hellerstein *et al.*, 1999; Aubert & Crépon, 2004; Dostie, 2006)

$$\log Y_{it} = \alpha \log L_{it}^A + \beta \log K_{it} \quad (1)$$

where: Y is the value added by firm i at time t , L^A is an aggregation of different types of workers, K is the capital stock, and μ is the error term.

The key variable in this production function is *the quality of labour aggregate* L^A . Let L_{ikt} be the number of workers of type k (women vs. men...) in firm i at time t , and μ be their productivity. We assume that workers of various types are substitutable with different marginal product. And each type of worker k is assumed to be an input in the production function. The aggregate can be specified as:

$$L_{it}^A = \sum_k \mu_{ik} L_{ikt} = \mu_{i0} L_{it} + \sum_{k>0} (\mu_{ik} - \mu_{i0}) L_{ikt} \quad (2)$$

where L_{it} is the total number of workers in the firm, μ_0 the productivity of the reference category of workers (*e.g.* men). Extensions of the basic methodology include enlarging the scope of workers' type, such as race and marital status, *e.g.* Hellerstein & Neumark (1995), Hellerstein *et al.* (1999) or age Vandenberghe & Waltenberg (2010). Here types refer exclusively to different gender or (as part of a extension aimed at assessing the robustness of our results) gender interacted with white- vs. blue-collar status.

If we further assume that a worker has the same marginal product across firms, we can drop subscript i and rewrite equation (2) as:

$$\log L_{it}^A = \log \mu_0 + \log L_{it} + \log (1 + \sum_{k>0} (\lambda_k - 1) P_{ikt}) \quad (3)$$

where $\lambda_k \equiv \mu_k/\mu_0$ the relative productivity of type k worker and $P_{ik} = L_{ik}/L_{i0}$ is the proportion/share of type k workers (e.g. share of women..) over the total number of workers in firm i .

Since $\log(1+x) \approx x$, we can approximate (3) by:

$$\text{Log } L_{it}^A = \log \mu_0 + \log L_{it} + \sum_{k>0} (\lambda_k - 1) P_{ikt} \quad (4)$$

And the production function becomes:

$$\log Y_{it} = \alpha [\log \mu_0 + \log L_{it} + \sum_{k>0} (\lambda_k - 1) P_{ikt}] + \beta \log K_{it} \quad (5)$$

Or, equivalently, if $k=0,1, \dots, N$ with $k=0$ being the reference group (e.g. men)

$$y_{it} = A + \alpha l_{it} + \eta_1 P_{i1t} + \dots + \eta_N P_{iNt} + \beta k_{it} \quad (6)$$

where:

$$\begin{aligned} A &= \alpha \log \lambda_0 \\ \lambda_k &= \mu_k/\mu_0 \quad k=1 \dots N \end{aligned}$$

$$\eta_1 = \alpha (\lambda_1 - 1)$$

....

$$\eta_N = \alpha (\lambda_N - 1)$$

$$y_{it} = \log Y_{it}$$

$$l_{it} = \log L_{it}$$

$$k_{it} = \log K_{it}$$

Note first that (6) being loglinear in P the coefficients can be directly interpreted as the percentage change in productivity of a 1 unit (here 100%) change of the considered type of workers' share among the employees of the firm. Note also that, strictly speaking, in order to obtain a type's relative productivity, (i.e. λ_k), coefficients η_k have to be divided by α , and 1 needs to be added to the result.

In order to test the null hypothesis of no gender wage discrimination we still need to define a labour costs/wage equation to obtain an estimate of the gender wage differential. Under the identifying assumptions of spot labour markets and cost-minimizing firms, male and female workers should be paid according to their marginal product. Let the total labour costs of a firm (LC) be decomposed in two components: labour costs with male workers ($k=0$) and labour costs with female workers ($k>0$).

By assumption, firms operate in the same labour market.¹¹ So they pay the same wages to the same category of workers (we can thus drop subscript i), which in our framework is the only feature that differentiates workers. Let π_k stand for the remuneration of type k workers. Then:

$$LC_{it} = \sum_k \pi_k L_{ikt} = \pi_0 L_{it} + \sum_{k>0} (\pi_k - \pi_0) L_{ikt} \quad (7)$$

Taking the log and using again $\log(1+x) \approx x$, we can approximate this by:

$$\log LC_{it} = \log \pi_0 + \log L_{it} + \sum_{k>0} (\Phi_k - 1) P_{ikt} \quad (8)$$

where the Greek letter $\Phi_k \equiv \pi_k / \pi_0$ denotes the yearly labour costs differential between women ($k>0$) and men ($k=0$), hereafter referred to as the gender wage differential, and $P_{ik} = L_{ik} / L_{i0}$ is the proportion/share of type k workers over the total number of workers in firm i .

The labour costs/wage model finally becomes:

$$w_{it} = B + \rho_1 P_{i1t} + \dots + \rho_N P_{iNt} \quad (9)$$

where:

$$\begin{aligned} B &= \ln \pi_0 \\ \Phi_k &\equiv \pi_k / \pi_0 \quad k=1, \dots, N \\ \rho_1 &= \Phi_1 - 1 \\ &\dots \\ \rho_N &= \Phi_N - 1 \\ w_{it} &= \ln LC_{it} - \ln L_{it} \end{aligned}$$

Note in particular that the dependent variable corresponds to the average labour costs per worker. By estimating equation (9) we can directly obtain an estimate of the gender wage differential by adding 1 to estimated ρ_k .

The gender wage discrimination test can now be easily formulated. Assuming spot labour markets and cost-minimizing firms the null hypothesis of no gender wage discrimination for type k worker implies $\lambda_k = \Phi_k$. Moreover, the gap between the gender productivity differential and the gender wage differential provides a quantitative measure of the extent of gender wage discrimination.¹² As it will be made clear in Section 5, this is a test we can easily implement in our econometric specifications

¹¹ At least at the sectoral level (NACE2). See next Section 3.2 below to see how we allow for sector (unobserved) specificities by resorting to fixed effects.

¹² We assume for presentational simplicity that women are less productive than men, so that the gender productivity differential is below 1.

of the production function and the labour costs equation.

Assuming that the LP polynomial is a good proxy for short- to medium term productivity shocks (an unobserved variable potentially correlated with gender mix if women are over represented among temp/part-time contacts), then the unaccounted part of the gender mix variance within firm — the one ultimately providing identification — probably reflects the overall rising propensity of women to work or to be allowed to in some sectors due to technical change/retirement of cohorts of men embodying outdated gender biased technological constraints. Table 1 in Section 4 shows that the overall share of women was on the rise over the period covered by our survey data.

3.2. Identifying the production function

We now consider the econometric version of our linearised Cobb-Douglas model (10). Note first that we have added a matrix F_{it} , wherein we concentrate region (#3), year (#8), sector¹³ (#76), and (the log of) average hours worked.¹⁴ The latter aims at capturing women's higher propensity to work part-time and controlling for spurious productivity and labour costs effects this may entail when the share of female work changes over time inside a firm.

The extension of the production function by introducing year, sector and region dummies allows for systematic and proportional productivity variation among firms along these dimensions. This assumption can be seen to expand the model by controlling for year- and sector- specific productivity shocks, labour quality and intensity of efficiency wages differentials across sectors and other sources of systematic productivity differentials (Hellerstein & Neumark, 1995). More importantly, since the data set we used did not contain sector price deflators, the introduction of these sets of dummies can control for asymmetric variation in the price of firms' outputs at sector. An extension along the same dimensions is made with respect to the labour costs equation. We recall that the labour costs equation is definitional: under the assumption of cost-minimizing firms that operate in the same competitive labour market, all workers in the same demographic categories earn the same wage. By introducing year, region and sector controls we consider the possibility that firms operate in year-, region- and sector-specific labour markets¹⁵ and, therefore, allow for wage variation along these dimensions. Of course, the assumption of segmented labour markets, implemented by adding linearly to the labour costs equation the set of dummies, is valid as long

¹³ NACE2 level. See Appendix for detailed list.

¹⁴ Total hours worked on an annual basis divided by the number of employees (part-time, full-time.).

¹⁵ It is probably the sector dimension that is the most relevant in the case of Belgium.

there is proportional variation in wages by gender along those dimensions (Hellerstein & Neumark, 1995).

$$y_{it} = A + \alpha l_{it} + \eta_1 P_{i1t} + \dots + \eta_N P_{iNt} + \beta k_{it} + \gamma F_{it} + \varepsilon_{it} \quad (10)$$

where $\varepsilon_{it} = \theta_i + \omega_{it} + \sigma_{it}$

where: $cov(\theta_i, P_{i1,t}) \neq 0$ and/or $cov(\theta_i, P_{i2,t}) \neq 0$, $cov(\omega_{it}, P_{i1,t}) \neq 0$ and/or $cov(\omega_{it}, P_{i2,t}) \neq 0$, $E(\sigma_{it})=0$

But from an econometric point of view, the main challenge consists of dealing with the various constituents of the residual ε_{it} of the production function. First, the *unobservable* (time-invariant) *heterogeneity* across firms, θ_i . The latter corresponds to specific characteristics of the firm, which are unobservable but driving the productivity while also being correlated with the explanatory variable of interest (here the share of women vs. men); for example the age of the plant, the vintage of capital used. Male workers might be overrepresented among plants built a long time ago, that use older heavy equipment that is intrinsically more difficult to operate for female employees. The panel structure of our data allows us to use fixed-effects or other within methods like first difference, attenuating that problem in many of the specifications.

However, the greatest econometric challenge is to go around *simultaneity or endogeneity bias* (Griliches & Mairesse, 1995). The economics underlying that concern is intuitive. In the short run firms could be confronted to productivity shocks, ω_{it} , (say, a positive shock due to a turnover, itself the consequence of a missed sales opportunity). Contrary to the econometrician, firms may know about this and respond by expanding recruitment of temporary- or part-time staff. Since the latter is predominantly female, we should expect that the share of female employment should increase in periods of positive productivity shocks and decrease in periods of negative shocks. This would generate spurious positive correlation between the share of female labour force and the productivity of firms, thereby leading to underestimated OLS estimates of the gender productivity differential.

Instrumenting the age by lagged values is a strategy regularly used in the production function literature (Arellano & Bond, 1991) to cope with this short-term simultaneity bias. Nevertheless, it has some limits, among which concerns about the quality of lagged values as instruments, and the large standard errors usually found, which make it difficult to draw solid conclusions.¹⁶ A development of that procedure, which has been proposed by Blundell & Bond (2000), is a system-

¹⁶ These limits have been acknowledged by Aubert & Crépon (2004), who applied such strategy to French data, and are also mentioned by Dostie (2006) or Roodman (2006).

GMM, in which the endogenous variables are instrumented with variables considered to be uncorrelated with the fixed effects and estimated by GMM. Still in this case, there are at least two types of problems: *i*) the estimated results are typically extremely sensitive to a great number of methodological choices (e.g., the number of lags for each variable), and, *ii*) instruments are often weakly identified, casting doubts on the quality of the estimations.

3.3. The intermediate input proxy approach to simultaneity bias

An alternative that seems to be particularly promising and relevant given the content of our data is to adopt the approach suggested by Levinsohn & Petrin (2003) and used, for example, by Dostie (2006). Their idea is that firms primarily respond to productivity shocks ω_{it} by adapting the volume of their *intermediate inputs*. Whenever such kind of information is available in a data set — which happens to be the case with ours as we have information on intermediate consumption (more on this in Section 4) — they can be used to proxy productivity shocks. An advantage with respect to the system-GMM method mentioned above is that this method based on intermediate inputs does not carry the burden of relying on instruments that lack a clear-cut economic meaning and which are, as mentioned above, typically weak.¹⁷ Moreover, by using the LP method, the number of discretionary methodological choices that have to be made by the researchers is reduced, contributing to providing results which are easier to understand and to compare with others in the literature.¹⁸

Formally, the demand for intermediate inputs would be a function of productivity shocks as well as the level of capital:

$$int_{it} = I(\omega_{it}, k_{it}) \quad (11)$$

Assuming this function is monotonic in ω and k , it can be inverted to deliver an expression of ω_{it} as a function of int and k . Expression (10) thus becomes:

$$y_{it} = A + \alpha l_{it} + \eta_1 P_{i1t} + \dots + \eta_N P_{iNt} + \beta k_{it} + \gamma F_{it} + \theta_i + \omega_{it}(int_{it}) + \varepsilon_{it} \quad (12)$$

with: $\omega_{it}(int_{it})$ that can be approximated by a polynomial expansion in int .

¹⁷ That is instruments are only weakly correlated with the included endogenous variables.

¹⁸ For example, employing the Arellano-Bond method, Aubert & Crépon (2004) have used a different number of lags for labour (2 lags) and other variables (all lags). Although they chose to reduce the number of lags for labour in order not to inflate too much the orthogonality conditions, it is not clear what procedure has been used to set those lags on the specific values they have chosen. We do not know whether their main results would be robust to different lag choices.

While the latter technique (in combination of firm fixed effects) is our preferred one, we have decided to report results of different econometric techniques, because of the well-known challenges and controversies involved in the estimation of any production function (Griliches & Mairesse, 1995).

Having identified our preferred econometric model, we can precise the source of identifying variance of both λ_k and Φ_k in equations (6),(9). It obviously comes from variation of the share of women. But could this reflect employer's preferences? ¹⁹ Neumark (1988) shows that if employer's discriminatory behavior concerns the share of female employment in each firm and if discrimination intensity of employers' is variable, then the variation of the share of female in each firm is the result of the variation in employer's discriminatory intensity. But our estimation uses within- rather than between-variation. The source of change at the firm-level in the share of female must come from elsewhere. Our source of identification cannot come from firm- specific "preferences" as to gender mix. These are wiped out by the fixed effects if we assume that they do not vary in the short- to medium run. What is more, assuming that the LP polynomial is a good proxy for short- to medium term productivity shocks (an unobserved variable potentially correlated with gender mix if women are over represented among temp/part-time contacts), then the unaccounted part of the gender mix variance within firm — the one ultimately providing identification here — is likely to reflect the overall rising propensity of women to work or to be allowed to in some sectors due to technical change (deindustrialisation) /retirement of cohorts of men embodying outdated gender biased technological constraints. The rising overall share of women in our sample (from 26 to 28 % between 1998 and 2006) is supportive of this assumption (Table 1).

¹⁹ In reference to Becker's (1957) taste-based discrimination theory or Arrow's (1972) theory of statistical discrimination.

4 Data and descriptive statistics

The firm-level data we use in this paper involves input and output variables of close to 9,000 firms of the Belgian private economy observed along the period 1998-2006. The data set matches financial and operational information retrieved from Belfirst with data on individual characteristics of all employees working in the firms, obtained from the Belgium's Social Security register (the so-called Carrefour database). The data set covers all sectors in the Belgian non-farming private economy, identified by NACE2 code⁶. Monetary values are expressed in nominal terms.

The productivity outcome corresponds to the firms' net value added: the value of output less the values of both intermediate consumption and consumption of fixed capital. The measure of labour costs, which was measured independently of net-value added (Figure 1), includes the value of all monetary compensations paid to the total labour force (both full- and part-time, permanent and temporary), including social security contributions paid by the employers, throughout the year. The summary statistics of the variables in the data set are presented in Table 1 and Table 2.

As we have mentioned in the previous section, we control for price variation in firms output by using a set of dummies for sector, year and their interaction. In our empirical analysis we use net value-added as the measure of firms' output. Capital input is measured by fixed tangible assets, while labour input corresponds to total number of employees, including both full- and part-time and under permanent and temporary contract, weighted by a measure (hours worked annually) of relative work intensity in the firm vis-à-vis the sample average.

The fact that we cannot distinguish part- from full-time workers and workers under permanent and temporary contract is an important limitation of our empirical analysis, since women are known to be overrepresented in part-time and temporary contract. However, note in Table 1 the presence of *average worked hours*. It is obtained by dividing the total number of hours in the firms (on an annual basis) by the number of employees (full-time or part-time ones indistinctively). We systematically include this ratio among our control variables. The reason for this is quite straightforward. There is evidence in Table 1 that average hours worked is negatively correlated with the share of female work. It fell from 1576 hours per employee in 1998 to 1517 hours in 2006 while the share of women rose from 26% to 28% over the same period of time. Lesser hours per employee — driven by a higher degree of feminisation of the workforce — logically reflects

women's higher propensity to work part-time. But this is also something that crucially needs to be controlled for, in order to properly capture the productivity (and labour costs) effect of changes in the share of female workers.

Table 1: Belfirst-Carrefour panel. Basic descriptive statistics. Mean (Standard deviation in italics).

Year	Nobs	Net value-add (th.€)	Labour costs (th.€)	Number of employees	Capital (th.€)	Average hours worked ^a	Share of female	Share of blue-collar female	Share of blue-collar male	Share of white-collar female	Share of white-collar male
1998	7584	7,760	4,800	108	6,388	1576	0.263	0.085	0.486	0.177	0.251
		<i>50,301</i>	<i>32,805</i>	<i>474</i>	<i>99,443</i>	<i>502</i>	<i>0.245</i>	<i>0.168</i>	<i>0.341</i>	<i>0.205</i>	<i>0.231</i>
1999	7743	8,192	5,017	111	6,548	1576	0.266	0.085	0.482	0.180	0.252
		<i>54,668</i>	<i>32,455</i>	<i>475</i>	<i>103,365</i>	<i>310</i>	<i>0.244</i>	<i>0.167</i>	<i>0.340</i>	<i>0.205</i>	<i>0.229</i>
2000	7929	8,837	5,314	114	6,857	1566	0.271	0.085	0.475	0.185	0.254
		<i>55,296</i>	<i>32,539</i>	<i>472</i>	<i>111,964</i>	<i>324</i>	<i>0.244</i>	<i>0.166</i>	<i>0.339</i>	<i>0.207</i>	<i>0.228</i>
2001	8121	9,027	5,646	121	7,477	1574	0.274	0.084	0.468	0.189	0.258
		<i>53,836</i>	<i>32,959</i>	<i>511</i>	<i>119,272</i>	<i>883</i>	<i>0.244</i>	<i>0.164</i>	<i>0.339</i>	<i>0.209</i>	<i>0.228</i>
2002	8262	9,565	6,172	128	8,043	1544	0.275	0.082	0.462	0.192	0.263
		<i>59,781</i>	<i>39,160</i>	<i>690</i>	<i>130,471</i>	<i>343</i>	<i>0.243</i>	<i>0.162</i>	<i>0.339</i>	<i>0.210</i>	<i>0.230</i>
2003	8353	10,128	6,384	127	8,508	1531	0.276	0.082	0.459	0.194	0.265
		<i>58,778</i>	<i>37,988</i>	<i>643</i>	<i>138,520</i>	<i>301</i>	<i>0.243</i>	<i>0.161</i>	<i>0.339</i>	<i>0.211</i>	<i>0.230</i>
2004	8355	10,954	6,667	129	8,870	1542	0.276	0.081	0.456	0.194	0.268
		<i>63,694</i>	<i>37,649</i>	<i>644</i>	<i>147,481</i>	<i>246</i>	<i>0.242</i>	<i>0.161</i>	<i>0.338</i>	<i>0.210</i>	<i>0.230</i>
2005	8338	11,438	6,912	132	8,052	1525	0.276	0.080	0.454	0.196	0.270
		<i>64,558</i>	<i>37,691</i>	<i>645</i>	<i>62,724</i>	<i>276</i>	<i>0.242</i>	<i>0.159</i>	<i>0.338</i>	<i>0.210</i>	<i>0.230</i>
2006	8261	12,367	7,311	134	8,250	1517	0.280	0.080	0.448	0.200	0.272
		<i>68,878</i>	<i>39,686</i>	<i>638</i>	<i>61,954</i>	<i>1666</i>	<i>0.242</i>	<i>0.158</i>	<i>0.336</i>	<i>0.212</i>	<i>0.230</i>

a: Total number of hours worked during the year divided by the total number of employee (full-time or part-time ones).

Table 2: *Belfirst-Carrefour panel. Basic descriptive statistics, pooled data*

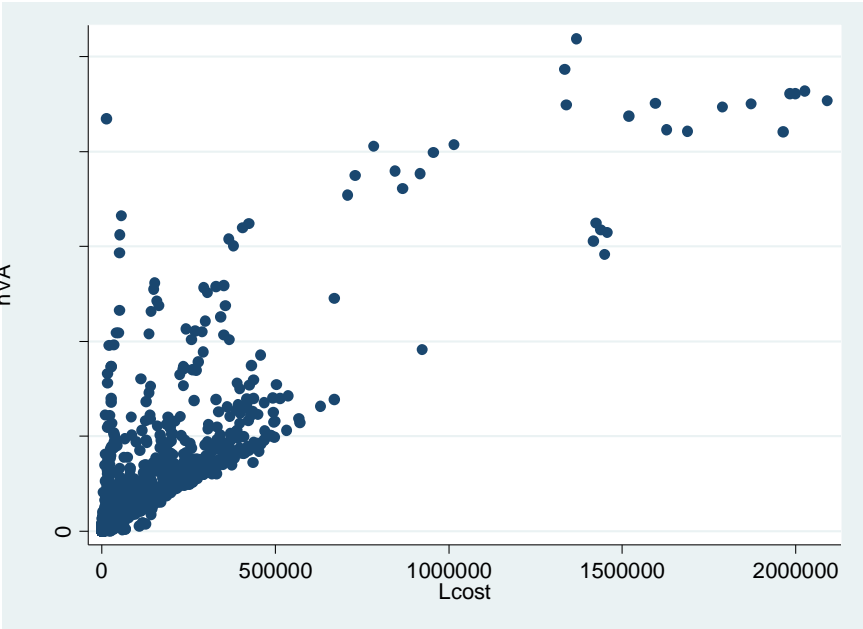
Firm size	Nobs
1-49	44354
50-99	14664
100+	13928
Region	
Brussels	10722
Vlaanderen	46008
Wallonia	16216

Figure 1 shows an expected pattern: a positive relation between firms' net value added (our measure of output) and their labour costs, with an overwhelming majority of firms reporting lower labour costs than their net value added.²⁰ Figure 2 reveals that productivity variance is higher than labour costs variance. In its lower panel, it also suggests that both average labour costs and productivity decline with the (rising) share of women employed by a firm.

Finally, intermediate inputs play a key role in our analysis, as they are central to our strategy to overcome the simultaneity bias. It is calculated here as the differences between the firm's turnover (in nominal terms) and its net value-added. It reflects the value of goods and services consumed or used up as inputs in production by enterprises, including raw materials, services and various other operating expenses.

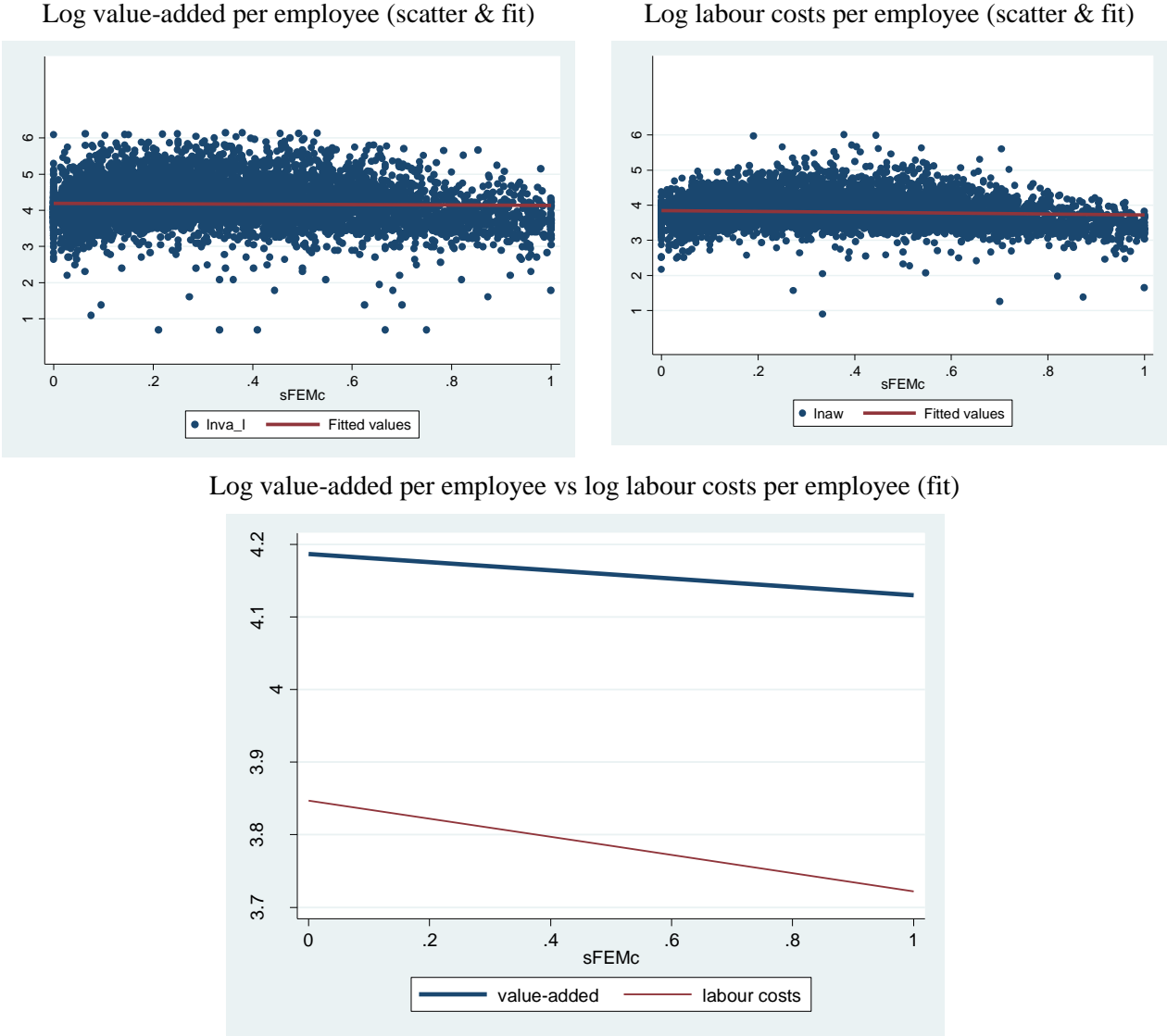
²⁰ The average productivity/labour costs ratio is 1.42.

Figure 1: Firms' labour costs versus firms' net value added (in th. €), pooled data



Source: Carrefour, Belfirst

Figure 2: Share of women in firms' workforce (on the horizontal axis) versus firms' i) log of net value added per employee ii) log of labour costs per employee. Year 2006. Scatter plot and linear fit



Source: Carrefour, Belfirst

5 Econometric Analysis

This section starts by complementing the description and justification of our methodological choices initiated in the previous section (subsection 5.1); next, it analyses the results of our estimations (subsection 5.2) and, finally, interprets the results in light of existing gender economic discrimination theories and previous evidence for the Belgian labour market (subsection 5.3).

5.1 Empirical Strategy

In Table 3 we present results of the independent estimation of production and the labour costs equations under six alternative econometric specifications: standard OLS using total variance [1] then OLS using only between-firm (or cross-sectional) variance [2]. Then comes the LP intermediate consumption “proxy” using total variance [3]. The next model uses first-differenced variables [4]. The fifth model is the within model (where each observation has been centred of the firm average over the duration of the panel). Finally, our preferred model is the one that combines the HP idea and the within-firm model [6].

Further ahead, in Table 4, we will focus on the simultaneous estimation of the production and labour costs functions using our preferred model [6] with the aim of assessing the statistical significance of the gap between gender productivity vs. labour costs differentials.

Specification [6] in Table 4 is a priori the best insofar as the coefficients of interest are identified from within-firm variation and that it controls for potential heterogeneity and simultaneity biases using LP’s intermediate input proxy strategy. Heterogeneity bias might be present since our sample covers all sectors of the Belgian private economy and the list of controls included in our models is limited. Even if the introduction of the set of dummies can account for most of this bias, the «within firm» transformation [5], [6] (or the first-differing one [4]) are still the most powerful way to account of inter-firm unobserved heterogeneity.

On the other hand, the endogeneity in input choice is a largely well documented problem in the production function estimation literature (*e.g.* Griliches and Mairesse, 1995) and also deserved to be properly treated. Moreover, given that our data do not distinguish between part- and full-time and temporary and permanent workers and that there is evidence from the Belgian labour market indicating that women tend to be overrepresented in part-time and temporary employment, the presence of simultaneity bias may underestimate the OLS estimates of the gender productivity differential.

Despite the considerations we made in the previous paragraphs, we believe specifications [1] to [4] provide valuable information about the presence and magnitude of biases, so that we will draw tentative evidence from comparison of the results of the alternative specifications.

We now make a final a justification for our preferred joint estimations of production and labour cost equations (Table 4). We recall that the focus of our analysis is the implementation of the gender wage discrimination test, which involves testing the equality of estimates of productivity (λ) and labour costs (Φ) differentials, obtained from estimations of the production function and the labour costs equations. Options here are essentially twofold.

First, joint estimation of the two equations (using *e.g.* the *SUREG*, Stata command). We recall that the arguments for joint estimation — what corresponds to *system FGLS* estimation in Wooldridge (2002)'s terminology²¹ — are essentially two. One is that joint estimation provides a direct way to implement a Wald test of the equality of a non-linear combination of coefficients across equations. If there are unobservables in both equations that bias the estimates of λ and Φ , as long as they affect the two equations equally, which should occur under the null, their effect on the Wald equality test is neutralized. Another is that joint estimation makes use of cross-equation correlations in the errors, thereby increasing the efficiency (i.e. generate smaller standard errors) of the coefficient estimates. Alternatively, one can perform so-called *system OLS* estimation. This consists of estimating the two equations separately, but to use those estimates to construct a cluster-adjusted²² robust sandwich variance-covariance matrix, which can be used to perform a Wald test of equality of the two coefficients.²³

The choice between system OLS and system FGLS can be viewed as a trade-off between robustness and efficiency. On the one hand, system OLS is more robust (i.e. generate coefficient that are less likely to be biased). It is consistent under the milder assumption of contemporaneous exogeneity, while the consistency of system FGLS is conditional on strict exogeneity of the regressors. Moreover, the Wald test computed from system OLS estimation can be made robust to arbitrary heteroskedasticity and serial correlation in the error term, while system FGLS does so under the assumption of system homoskedasticity. In principle, we could construct a cluster-adjusted robust sandwich variance-covariance matrix from the FGLS estimates. However, the Stata command that implements FGLS, *SUREG*, does not permit its computation from standard commands. On the other hand, system FGLS takes advantage of increased efficiency from cross-equation correlations in the errors.

²¹ See chapter 7 of Wooldridge (2002) for a derivation of the properties of system OLS and system FGLS estimators.

²² Here, a cluster is a firm.

²³ See Weesie (2000) for a description of the Stata procedure that constructs a cluster-adjusted robust sandwich estimator from two or more sets of independent estimates.

We decided to implement *system OLS* in addition to the more common *system FGLS* (used for instance by Hellerstein &Neumark (1995) and Hellerstein *et al.* (1999) for four reasons. First, because we are using panel data, so that the error term should normally be serially correlated for the same firm, the ability to control for arbitrary heteroskedasticity and serial correlation across time is a strong advantage. Second, the advantage of controlling for potential unobservables is substantially smaller in our case: while Hellerstein &Neumark (1995) and Hellerstein *et al.* (1999) used cross section data and implemented standard OLS and IV estimators, instead, we use panel data and implement estimation procedures specifically designed to deal with potential biases due to unobservables. Third, the importance of cross-equation correlation in the errors needs to be assessed vis-à-vis the efficiency of the estimates obtained from independent estimations. In our case, the precision of coefficient estimates using system OLS is fairly satisfactory. Fourth and last, the assumption of strict exogeneity is very strong for production function estimation. That said, the efficiency gains associated with system FGLS seem to be high for our data set: the cross-equation correlation of the residuals is high both for the raw and the transformed data, respectively 69%, for total-firm variation, and 56% for within-firm variation, and 60%, for total-firm variation, and 40% for within-firm variation.

5.2 Empirical Results

Table displays the parameter estimates of the production and labour costs functions when these are estimated separately. Reported coefficients in the upper parts of the table correspond to $\eta = \alpha(\lambda-1)$; $\rho = \Phi - 1$ in equations 6 & 9.

The lower part of Table 3 contains the estimates of the gender productivity (λ) and labour costs (Φ) differentials. Estimated λ point at lower productivity inside firms employing more women. Male to female productivity differentials range for 0 to -18 percentage points. Those for Φ are significant and point negative labour costs differentials for women. These range from 0 to -17 percentage points.

The crucial issue, however, is the gap between these gender differentials as it captures the intensity of gender wage discrimination. We report different estimates of this gap on the bottom line of Table 3. OLS estimates (column [1]) suggest that women in the Belgian labour market are paid 12 percentage point less than what their (relative) productivity would imply. Turning to the between-

firm estimates (were we solely use the between firm variance), we get an even larger gap of 13 percentage points. But focusing on the within-firm variance (in order to account for time-invariant unobserved heterogeneity) considerably reduces that gap. Indeed, estimates reported in column [5] translate into a now negative gap of about 3 percentage points. And when we combine the within approach (to control for time-invariant heterogeneity) and the LP's proxy strategy to control for short-term endogeneity, we get a negative gap of 6 percentage points. In other words, the gender labour costs differential is smaller than the productivity differential. Although these results require further qualifications (more on this below), they suggest that most of the evidence in support of gender pay discrimination vanishes once cross-firm unobserved heterogeneity and simultaneity bias have been controlled for.

The dramatic reduction of the differential gap when moving from total- to within-firm variance constitutes important evidence in support of controlling for cross-firm heterogeneity and rejecting OLS [1], between [2] on LP-only [3] estimates. This is particularly true for the labour costs equation. The within-firm labour costs differential is much smaller (6 percentage points [5], [6]) than in previous models (17 percentage points with OLS [1])²⁴ see lower part of Table 3).

The different estimates of the productivity differentials are also affected by the within transformation, although to a lesser extent than labour cost differentials. Controlling for unobserved heterogeneity and simultaneity bias combining within and LP [6] leads to gender productivity differentials of greater magnitude (-5 percentage points with OLS [1] vs. -13 percentage points with our preferred estimate [6], see lower part of Table 3).

The latter results accords with our initial prediction. Based on evidence for the Belgian labour market summarized in Meulders & Sissoko (2002), we were convinced that, if anything, the presence of simultaneity bias would lead to an underestimation of the gender productivity differential in OLS estimations. Our reasoning was the following: since in Belgium temporary

²⁴ Note that this estimate of the “gross” gender labour costs differential is quantitatively similar to previous studies of the gender wage differential in the Belgium labour market using individual level-data, wage equations and Oaxaca-Blinder decomposition methods. Jepsen (2001), using 1994-95 data from the ECHP (European Community Household Survey), finds an unadjusted wage gap ratio of 85%, which lowers to 83%, when part-time workers are included. For the same period, a report by the Belgian Federal Ministry of Employment and Labour, cited in Meulders & Sissoko (2002), using the same data set as Jepsen (2001) and another data set, SES (Structure of Earnings Survey), that only includes data for the private sector, finds an unadjusted gender pay gap of 16% in the private sector.

contract employment is asymmetrically concentrated in female employment,²⁵ we should expect that, if temporary employment is one, or the main, labour adjustment variable to shocks in firms economic environments, the share of female employment should increase in periods of positive productivity shocks and decrease in periods of negative productivity shocks. This would generate positive correlation between the share of female labour force and the productivity of firms, thereby leading to underestimated OLS estimates of the gender productivity differential. As we have just argued our results do confirm this prediction.

But strictly speaking, we cannot conclude to the absence of gender discrimination without properly *testing* for the equality of the gender productivity (λ) and labour costs differentials (Φ). Table 4 presents estimates of λ and Φ obtained from both system FGLS and system OLS estimations of the production function and the labour costs equation, and the p-values of Wald equality tests of these coefficients.

With system FGLS, the estimates of λ and Φ (and the resulting gaps) are approximately the same as those obtained from system OLS estimates (Table 4) and, as expected, the precision of the estimates increased slightly owing to the high correlation in the residuals across equations (around 60% for total-firm estimations and around 40%, for within-firm estimations). But in both cases high p-values of the Wald equality tests statistic (0.84 and 0.28 respectively) lead to the acceptance of the null hypothesis of no gender wage discrimination.

We have undertaken two further steps in our analysis to assess the robustness of these results. First, we have examined whether our results change much when we partition the sample in terms of *firm size*. Second, we go beyond the simple distinction between men and women and consider *the interaction of status* (blue-collar/white collar) *and gender*. Referring to equations 6 and 9, this means estimating these models with $k=0,1,2,3$ categories of workers, where the reference category in our case ($k=0$) are the blue-collar men. Note in particular that the white vs. blue-collar workers comparison is a way to somehow compensate for the lack of information on the level of education (which is one shortcoming of our data). For each of these extensions, the focus will be on the results of the model with intermediate inputs *à-la*-LP with firm fixed effects (exploiting within-firm variance). We also resort to both system FGLS (Table 5, panel A) and system OLS (Table 5, panel B) to assess the null hypothesis of no gender wage discrimination ($\lambda = \Phi$).

²⁵ The same could be said of part-time employment, but remember that we explicitly control for the latter by including average hours worked per employee (part-time or full-time employees confounded) in all our estimations.

The main results from these breakdowns do not differ in qualitative terms from those obtained using the overall sample. Whatever the method used (system FGLS or system OLS), we conclude to the absence of systematic gender discrimination when consider the breakdown according to white- vs. blue-collar status. Female workers get paid in relative terms slightly more than their relative productivity, which leads to the negative gaps reported in Table 5.A and 5.B. Yet, these are generally not statistically significant. It if only in large firms (100+) that we find evidence supportive of gender discrimination. Our system OLS estimate suggest a positive gap of about 6 percentage point, though the coefficient is not statistically significant (*i.e.* productivity higher than labour costs for women). System FGLS delivers a positive gap of 15 percentage points that is statistically significant, but only at the 1% level.

Table 3: Separate estimation of Production Function and Labour Costs Equation

Method:	1-OLS	2-Between	3-Intermediate inputs (Levinsohn-Petrin)	4-First-Differences	5-Within (firm fixed effects)	6-Within (firm fixed effects+ intermediate inputs LP)
Productivity equation						
Share Women	-0.045***	0.014	-0.021*	-0.068*	-0.072**	-0.103***
<i>p-value</i>	0.0000	0.4897	0.0348	0.0163	0.0025	0.0002
Controls	capital. number of employees. hours worked per employee + fixed effects: year. nace1. region	capital. number of employees. hours worked per employee + fixed effects: year. nace1. region	capital. number of employees. hours worked per employee + fixed effects: firm	capital. number of employees. hours worked per employee + fixed effects: firm	capital. number of employees. hours worked per employee + fixed effects: firm	capital. number of employees. hours worked per employee + fixed effects: firm
Nobs.	59 980	59 980	49 582	49 395	59 980	49 575
Labour-cost equation						
Share Women	-0.171***	-0.117***	-0.131***	-0.013	-0.063***	-0.065***
<i>p-value</i>	0.0000	0.0000	0.0000	0.3814	0.0000	0.0000
Controls	hours worked per employee+ fixed effects: year. nace1. region	hours worked per employee+ fixed effects: year. nace1. region	hours worked per employee+ fixed effects: year. nace1. region	fixed effects: firm. year	fixed effects: firm. year	fixed effects: firm. year
Nobs.	60 713	60 713	49 581	50 110	60 713	49 581
Productivity vs labour cost differentials						
Productivity diff. (λ)	0.95	1.02	0.98	0.90	0.91	0.87
Labour costs diff. (Φ)	0.83	0.88	0.87	0.99	0.94	0.94
Gap ($\lambda-\Phi$)	0.12	0.13	0.11	-0.09	-0.03	-0.06

*p < 0.05, **p < 0.01, *** p < 0.001

Table 4: Joint estimates of productivity and labour costs differentials. Within (firm fixed effects) + intermediate inputs (Levinsohn-Petrin). Cluster-robust estimation of standard-errors.

	Production diff. (λ): ref=men	Labour-cost diff (Φ): ref=men	Gap ($\lambda-\Phi$)	Wald Hyp. Test ($\lambda=\Phi$)	
				χ^2	Prob> χ^2
System FGLS	0.936	0.941	-0.005	0.04	0.8473
System OLS	0.881	0.941	-0.060	1.14	0.2863

*p < 0.05, **p < 0.01, *** p < 0.001

a: Simultaneous estimation accounting for possible correlation between residuals

b: Equations are estimated separately

Table 5: Joint estimates of productivity and labour costs differentials. Breakdown by firm size and labour market status (p-values in italics). Within (firm fixed effects)+ intermediate inputs (Levinsohn-Petrin). Cluster-robust estimation of standard-errors

A System FGLS^a

	Production diff. (λ): ref=men	Labour-cost diff (Φ): ref=men	Gap ($\lambda-\Phi$)	Wald Hyp. Test ($\lambda=\Phi$)	
				χ^2	Prob> χ^2
System FGLS*					
Firm size	ref=men	ref=men			
1-49	0.86	0.91	-0.046	1.84	<i>0.1744</i>
50-99	0.96	0.93	0.029	0.26	<i>0.6134</i>
>=100	1.21	1.06	0.151*	5.47	<i>0.0193</i>
Gender/Status	ref=blue-collar men	ref=blue-collar men			
blue-collar women	0.84	0.88	-0.041	0.97	0.3246
white-collar women	1.20	1.23	-0.025	0.65	0.4186
white-collar men	1.35	1.41	-0.056*	4.33	0.0374

*p < 0.05, **p < 0.01, *** p < 0.001

B System OLS^b

System OLS	Production diff. (λ): ref=men	Labour-cost diff (Φ): ref=men	Gap ($\lambda-\Phi$)	Wald Hyp. Test ($\lambda=\Phi$)	
				χ^2	Prob> χ^2
Firm size	ref=men	ref=men			
1-49	0.75	0.91	-0.154*	4.71	0.0300
50-99	0.86	0.93	-0.071	0.36	0.5459
>=100	1.12	1.06	0.059	0.21	0.6483
Gender/Status	ref=blue-collar men	ref=blue-collar men			
blue-collar women	0.80	0.83	-0.026	0.61	0.4356
white-collar women	0.96	1.16	-0.202	2.53	0.1120
white-collar men	1.09	1.32	-0.231	2.22	0.1366

*p < 0.05, **p < 0.01, *** p < 0.001

a: Simultaneous estimation accounting for possible correlation between residuals

b: Equations are estimated separately, but the estimates are used to construct a cluster-adjusted robust sandwich variance-covariance matrix.

c: See appendix for a presentation of NACE2 codes corresponding to these categories

5.3 Interpretation of Results

In interpreting the above empirical results it is helpful to bear in mind the benchmark definition of gender wage discrimination presented in Section 2.1: identifying market-wide and statistically significant gaps between gender productivity differentials and gender wage differentials. Recall that Hellerstein & Neumark (1995) empirical methodology does not provide a direct test of any particular theory of gender wage discrimination, rather, it supplies an empirical measure of the above benchmark concept of gender wage discrimination.

Nevertheless, although the Hellerstein & Neumark methodology does not provide a direct test for any particular theory of gender wage discrimination, we can still check which theories of gender wage discrimination are consistent with our empirical findings. Our core findings based on within-firm variation and the various extensions we carried out considering both firm- or worker traits (*i.e.* size and blue- or white-collar status) indicate that the null hypothesis of no gender wage discrimination holds. Indeed, although our results indicate that male and female labour do not provide the same services in the each firm, insofar as women, as a group, are significantly less productive than men, they do not reject the hypothesis that women get paid according to their lower productivity with respect to men.

6 Conclusion

In this paper we used firm-level data from a matched employer-employee data set to test for the presence of gender wage discrimination in the Belgian labour market. We identified gender wage discrimination from within-firm variation and used Levinsohn and Petrin (2003) structural production function estimator to control for the endogeneity in input choice. Our findings indicate that, on average, women earn 6% less than men but also that they are collectively 6-12% less productive than men.

The results of the implementation of the Wald test of equality of the gender wage differential and the gender productivity differential — or of the statistical significance of productivity-to-wage gap, ranging from 0 to -6 percentage points — lead us to the non-rejection of the null hypothesis that, under the assumptions of spot labour markets and cost-minimizing firms, women are not systematically discriminated against in earnings in the Belgian labour market.

In essence, these findings are consistent with the prediction of Becker (1957) that they are efficiency costs associated with gender-biased preferences by employers, and that competition should eliminate wage discrimination in the long run. The estimates of the gender labour costs differential we obtained also accord with those obtained in empirical studies using Oaxaca-Blinder decompositions based on wage equations to explain the sources of gender differences in pay in the Belgian labour market (Rycx & Tojerow, 2002). More importantly, due to the ability of Hellerstein & Neumark's methodology to supply a direct test for the gender wage discrimination hypothesis, we contribute with new evidence to the research programme dedicated to explaining the sources of the gender pay gap. Because we use firm-level data we are indeed able to estimate gender *productivity differences* alongside the traditional gender wage/labour costs differences, and show that the two are approximately aligned.

References

- Aigner, D. J. and Cain, G. G. (1977), 'Statistical Theories of Discrimination in Labor Markets', *Industrial and Labor Relations Review*, vol. 30, pp. 175-187.
- Altonji, J. G. and Blank, R. (1999), 'Race and Gender in the Labor Market'. In: *Handbook of Labor Economics*, Ed. O. Ashenfelter and D. Card, vol. 3C, Chapter 48, Amsterdam: North-Holland.

- Arellano, M. and S. Bond (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies*, 58, pp. 277-297.
- Arrow, K. (1972), *Models of Job Discrimination*. In: *Racial Discrimination in Economic Life*, Ed. Anthony Pascal, Lexington, MA: Lexington Books.
- Aubert, P. and Crépon, B. (2004), 'Age, salaire et productivité: la productivité des salariés décline-t-elle en fin de carrière?', *Économie et Statistiques*, No 368, pp. 43-63.
- Beblo, M., Beninger, D., Heinze A. and Laisney, F. (2003), *Methodological Issues Related to the Analysis of Gender Gaps in Employment, Earnings and Career Progression*, Final Report for the European Commission Employment and Social Affairs DG, Brussels: European Commission.
- Becker, G. (1957), *The Economics of Discrimination*, Chicago: University of Chicago Press.
- Bergmann, B. R. (1971), 'The Effect on White Incomes of Discrimination in Employment', *Journal of Political Economy*, 79(2), pp. 294-313.
- Bergmann, B. R. (1974), 'Occupational Segregation, Wages, and Profits when Employers Discriminate by Race or Sex', *Eastern Economic Journal*, 1, pp. 561-573.
- Blau, F. D. and Kahn, L. M. (2000), 'Gender Differences in Pay', *Journal of Economic Perspectives*, 14(4), pp. 75-99.
- Blinder, A. (1973), 'Wage Discrimination: Reduced Form and Structural Variables', *Journal of Human Resources*, 8(4), pp. 436-465.
- Blundell, R. and Bond, S. (2000), 'GMM Estimation with Persistent Panel Data: An Application to Production Functions', *Econometric Reviews*, 19(3), pp. 321-340.
- Borjas, G. J. and Bronars, S. G. (1989), 'Consumer Discrimination and Self-employment', *Journal of Political Economy*, 97(3), pp. 581-606.
- Bowlus, A. J. and Eckstein, Z. (2002), 'Discrimination and Skill Differences in an Equilibrium Search Model', *International Economic Review*, 43(4), pp. 1309-1344.
- Cain, G. C. (1986), 'The Economic Analysis of Labor Market Discrimination: A Survey'. In: *Handbook of Labor Economics*, Edited by O. Ashenfelter and R. Layard, vol. 1, Chapter 13, Amsterdam: North-Holland.
- Coate, S. and Loury, G. (1993b), 'Will Affirmative-Action Policies Eliminate Negative Stereotypes', *American Economic Review*, 83(5), pp. 1220-1240.

Crépon, B., N. Deniau, et S. Pérez-Duarte (2002). "Wages, Productivity, and Worker Characteristics: A French Perspective.", *Serie des Documents de Travail du CREST*, Institut National de la Statistique et des Etudes Economiques.

Dostie, B. (2006), Wages, Productivity and Aging, *IZA Discussion Paper*, No 2496, Bonn, Germany.

European Commission (2007), '*Tackling the Pay Gap between Women and Men*', Communication from the European Commission, COM (2007), 424 final.

Griliches, Z. and J. Mairesse (1995), Production functions: the search for identification. *NBER working paper*, No 5067, NBER, Ma.

Gunderson, M. (2006), 'Viewpoint: Male-Female Wage Differentials: How Can That Be?', *Canadian Journal of Economics*, 39(1), pp. 1-21.

Heckman, J. (1998), 'Detecting Discrimination', *Journal of Economic Perspectives*, 12(2), pp. 101-116.

Hellerstein, J. K. and Neumark, D. (1995), 'Sex, Wages and Productivity: An Empirical Analysis of Israel Firm-level Data', *International Economic Review*, 40(1), pp. 95-123.

Hellerstein, J.; Neumark, D.; Troske, K. (1999), Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations, *Journal of Labor Economics*, 17(3), pp. 409-446.

Institut pour l'égalité des Femmes et des Hommes (2006), *Femmes et hommes en Belgique. Statistiques et indicateurs de genre. Edition 2006*, Bruxelles.

Jepsen, M. (2001), 'Evaluation des différentiels salariaux en Belgique: homme-femme et temps-partiel-temps plein', *Reflète et Perspectives de la Vie Économique*, 40(1), pp. 91-99.

Levinsohn, J. and A. Petrin (2003), Estimating production functions using inputs to control for unobservables, *Review of Economic Studies*, 70 (2), 317-341

Meulders and Sissoko (2002), *The Gender Pay Gap in Belgium*, Report to the Expert Group on Gender and Employment, Brussels: Department of Applied Economics of Free University of Brussels (DULBEA-ETE).

Neumark, D. (1988), 'Employer's Discriminatory Behavior and the Estimation of Wage Discrimination', *The Journal of Human Resources*, 23(3), pp. 279-295.

- Neumark, D. (1996), 'Sex Discrimination in Restaurant Hiring: An Audit Study', *The Quarterly Journal of Economics*, 111(6), pp. 915-941.
- Oaxaca, R. (1973), 'Male-female Wage Differentials in Urban Labor Markets', *International Economic Review*, 14, pp. 693-709.
- Phelps, E. (1972), 'The Statistical Theory of Racism and Sexism', *American Economic Review*, 4, pp. 659-661.
- Roodman, D. (2006). "How to Do xtabond2", North American Stata Users' Group Meetings 2006 8, Stata Users Group.
- Rosén, A. (2003), 'Search, Bargaining and Employer Discrimination', *Journal of Labor Economics*, 21(4), pp. 807-828.
- Rycx, F. & I. Tojerow (2002), "Inter-industry Wage Differentials and the Gender Wage Gap in Belgium," *Brussels Economic Review/Cahiers Economiques de Bruxelles*, Editions du DULBEA, ULC, Department of Applied Economics (DULBEA), 45(2), pp. 119-141.
- Sasaki, M. (1999), 'An Equilibrium Search Model with Co-worker Discrimination', *Journal of Labor Economics*, 17(2), pp. 377-407.
- Vandenberghe, V. and F.D. Waltenberg (2010), Ageing Workforce, Productivity and Labour costs of Belgian Firms, *IRES WP*, No 2010-3, Economics School of Louvain, UCL, Louvain-la-Neuve.
- Weesie, J. (2000), 'Seemingly Unrelated Estimation and the Cluster-adjusted Sandwich Estimator', *Stata Technical Bulletin*, 9(52), pp. 34-46.
- Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Mass.: MIT Press.

Annex : Sectors (Industry, Commerce and Service) and NACE2 codes/definitions

- 10="I_Industries alimentaires "
- 11="I_Fabrication de boissons"
- 12="I_Fabrication de produits à base de tabac"
- 13="I_Fabrication de textiles"
- 14="I_Industrie de l'habillement"
- 15="I_Industrie du cuir et de la chaussure"
- 16="I_Travail du bois et fabrication d'articles en bois et en liège, à l'exception des meubles; fabrication d'articles en vannerie et sparterie"
- 17="I_Industrie du papier et du carton"
- 18="I_Imprimerie et reproduction d'enregistrements"
- 19="I_Cokéfaction et raffinage"
- 20="I_Industrie chimique"
- 21="I_Industrie pharmaceutique"
- 22="I_Fabrication de produits en caoutchouc et en plastique"
- 23="I_Fabrication d'autres produits minéraux non métalliques"
- 24="I_Métallurgie"
- 25="I_Fabrication de produits métalliques, à l'exception des machines et des équipements"
- 26="I_Fabrication de produits informatiques, électroniques et optiques"
- 27="I_Fabrication d'équipements électriques"
- 28="I_Fabrication de machines et d'équipements n.c.a."
- 29="I_Construction et assemblage de véhicules automobiles, de remorques et de semi-remorques"
- 30="I_Fabrication d'autres matériels de transport"
- 31="I_Fabrication de meubles"
- 32="I_Autres industries manufacturières"
- 33="I_Réparation et installation de machines et d'équipements"
- 35="I_Production et distribution d'électricité, de gaz, de vapeur et d'air conditionné"
- 36="I_Captage, traitement et distribution d'eau"
- 37="I_Collecte et traitement des eaux usées"
- 38="I_Collecte, traitement et élimination des déchets; récupération"
- 39="I_Dépollution et autres services de gestion des déchets"
- 41="I_Construction de bâtiments; promotion immobilière"
- 42="I_Génie civil"
- 43="I_Travaux de construction spécialisés"
-
- 45="C_Commerce de gros et de détail et réparation véhicules automobiles et de motocycles"
- 46="C_Commerce de gros, à l'exception des véhicules automobiles et des motocycles"
- 47="C_Commerce de détail, à l'exception des véhicules automobiles et des motocycles"
-
- 49="S_Transports terrestres et transport par conduites"
- 50="S_Transports par eau"
- 51="S_Transports aériens"
- 52="S_Entreposage et services auxiliaires des transports"
- 53="S_Activités de poste et de courrier"
- 55="S_Hébergement"
- 56="S_Restauration"
- 58="S_Édition"

59="S_Production de films cinématographiques, de vidéo et de programmes de télévision; enregistrement sonore et édition musicale"
60="S_Programmation et diffusion de programmes de radio et de télévision"
61="S_Télécommunications"
62="S_Programmation, conseil et autres activités informatiques"
63="S_Services d'information"
64="S_Activités des services financiers, hors assurance et caisses de retraite"
65="S_Assurance, réassurance et caisses de retraite, à l'exclusion des assurances sociales obligatoires"
66="S_Activités auxiliaires de services financiers et d'assurance"
68="S_Activités immobilières"
69="S_Activités juridiques et comptables"
70="S_Activités des sièges sociaux; conseil de gestion"
71="S_Activités d'architecture et d'ingénierie; activités de contrôle et analyses techniques"
72="S_Recherche-développement scientifique"
73="S_Publicité et études de marché"
74="S_Autres activités spécialisées, scientifiques et techniques"
75="S_Activités vétérinaires"
77="S_Activités de location et location-bail"
78="S_Activités liées à l'emploi"
79="S_Activités des agences de voyage, voyagistes, services de réservation et activités connexes"
80="S_Enquêtes et sécurité"
81="S_Services relatifs aux bâtiments; aménagement paysager"
82="S_Services administratifs de bureau et autres activités de soutien aux entreprises"
92="S_Organisation de jeux de hasard et d'argent"
93="S_Activités sportives, récréatives et de loisirs"
94="S_Activités des organisations associatives"
95="S_Réparation d'ordinateurs et de biens personnels et domestiques"
96="S_Autres services personnels"
97="S_Activités des ménages en tant qu'employeurs de personnel domestique"
98="S_Activités indifférenciées des ménages en tant que producteurs de biens et services pour usage propre"
99="S_Activités des organisations et organismes extraterritoriaux"

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

