UNIVERSITY OF CALIFORNIA, BERKELEY
ECONOMICS DEPARTMENT

FIRM BEHAVIOR AND IMMIGRANT LABOR PRODUCTIVITY

Raymundo M. Campos-Vazquez

Preliminary version
Comments welcome
rcampos@econ.berkeley.edu

June 29, 2007

ABSTRACT

The goal of the paper is the estimation of the marginal productivity of immigrants relative to natives. One reason why firms may hire immigrants is because immigrants are more productive than natives. I test this hyphotesis using a confidential matched employer-employee dataset from Germany for the years 1996-2004. Using a production function approach, I find that immigrants are similar to natives in terms of productivity and wages. Hence, the benefits of immigration need to come from non-wage sources or immigrant labor flexibility.

1 Introduction

Immigrant minorities, generally low skilled, are present in most of the developed world. As the net benefits of immigration for the whole economy and especially for low skilled workers are uncertain, immigration is a highly debated topic. Scholars have tried to disentangle the effects of immigration on natives' labor market outcomes. As in the political arena, economists have disagreed about these effects. Although there is abundant research on how immigrants affect natives at the aggregate or local level, research on how individual firms use immigrants is scant. In order to understand the net benefits of immigration, we first need to understand why there is demand for immigrant labor. The fastest answer to this question is that immigrants are paid a lower wage which causes a decline in the wages of workers, especially low skilled. A different story could be that immigrant labor provides an added value to the plant that hires immigrant labor. In this paper, I analyze whether this hypothesis is correct for a sample of firms from Germany.

Newer datasets, like matched employer-employee, can be used to estimate how immigrants and natives are used at the plant level. Analysis at the plant level can provide a clearer picture of how immigrants affect natives' labor market outcomes. For example, it can tell us what type of firms hire immigrants, how do they use them in the plant, the level of segregation of immigrants across plants and whether plants hire immigrants with a lower wage than natives. Moreover, it is possible to analyze the degree of substitutability between immigrants and natives across and within plants. Hence, as opposed to previous research, matched employer-employee data provides an opportunity to explain the link between immigration and natives' labor market outcomes.

The debate on the benefits of immigration is centered on the extent on how low-skilled immigrants put a downward pressure on native wages and employment outcomes. Previous research has relied on testing the degree of this pressure *assuming* immigrants and natives are perfect substitutes within some specified labor group aggregation.[1] This argument assumes natives and immigrants are equally productive within that labor group aggregation. Hence, as natives and immigrants are equally productive, the employer's decision to hire immigrants relies only on the wage margin. In other words, immigrants and natives are *assumed* to be similar and as such they produce the same amount of output in equal circumstances. However, one of the issues that has not been considered in the literature of immigration is precisely testing the assumption that productivity of immigrants is the same to natives' productivity.

In order to test this hypothesis, I use a strategy first proposed by Hellerstein et al. (1999).

[1] See Card (2001), Borjas (1999), Borjas (2003) and Ottaviano and Peri (2006).

Although they are interested in how women are underpaid relative to similar men, we are more interested in the relative marginal productivity of immigrants. Hence, the goal of the paper is the estimation of the marginal productivity of immigrants relative to natives. Once the marginal productivity of immigrants is estimated, it is contrasted to their relative wage. If immigrants are equally productive to natives but they earn less on average than natives, it is possible to conclude that firms are taking advantage of cheap labor. On the other hand, if we find evidence that immigrants are more productive than natives and at the same time immigrants are underpaid, we cannot conclude that immigrants are only hired because of a lower wage, but possibly also due to their higher productivity or another type of skills on attitudes that are valued by firms.

Why could we expect a different productivity among natives and immigrants? The literature on the self-selection of immigrants concludes that selection on unobservables is an important determinant for the decision to immigrate in the first place (Butcher, 1994). If workers are selected on unobservables, like their motivation to succeed, it is possible that immigrants put more effort in their work. Another possible channel relies in a power purchasing story. Immigrants think of their wage as goods they can purchase in their home country, given that goods are relatively cheaper there they put more effort than natives. If either hypothesis is correct, immigration could have more benefits than previously thought.

I use a unique longitudinal confidential data from Germany. This dataset is ideal to test differences in marginal productivities and wages. For a national representative sample of plants each year, I observe characteristics of the full workforce as well as the wage of each worker. This allows me to compare directly marginal productivity of immigrants relative to natives as well as their wages at the plant level for all workers in the plant. My results indicate that immigrants are as productive as natives and that they earn similar to slightly lower wages than natives. This finding is robust to several specifications. These results are not very informative about why firms hire immigrants. It is possible that there are other reasons of why firms hire immigrants without relying on differences in productivity. For example, one hypothesis that I am unable to test is whether immigrant labor causes a decrease in non-fringe benefits paid by the firm.

Germany is an interesting case to study the effects of immigration.[2] Germany started to recruit foreign workers as Guest Workers after the Second World War and stopped the recruiting process after the oil shock in 1973. From that moment it changed the immigration policy towards family reunification. Given the history of the country, Germany also had an

[2]Herbert (1990) describes the history of foreign labor in Germany during the twentieth century. Göktürk et al., eds (2006) and Chin (2007) describe a cultural history of the Guest Workers in Germany after the Second World War.

open border policy towards refugees until the beginning of the 1990s. Both processes caused a change in the cultural landscape, and transformed the country into a de facto diverse society when the Basic Law Act was modified in order to give German nationality to those born in German soil. In particular, the share of immigrants in the population is similar to the U.S. around 10 percent. Moreover, minorities among the immigrants can be easily identified, as in the United States. Turks, a group that is considered mostly low skilled, represent around 35 percent of immigrants in Germany, a similar figure arises for Mexicans in the U.S. The rest is formed by immigrants from former Yugoslavia (low skilled) and countries from the European Union.

The similitude between Germany and the U.S. does not restrict only to the characteristics of the immigrants and to the politics of immigration. Researchers have also found mixed effects of immigration on natives' labor market outcomes, but most of the empirical estimates suggest a close to zero effect of immigration. For the period 1985-1989, Pischke and Velling (1997) , using local labor market information, find that immigration does not incur in displacement effects. For a more long run perspective, Bonin (2005) uses the 1975-1997 period and replicates the analysis of Borjas (2003) at the aggregate level dividing immigrants and natives in experience and education cells. He finds that immigration does not have negative consequences on employment outcomes and at most a 10 percent increase in immigration will decrease wages by 1 percent. On the other hand, Glitz (2006) uses a quasi-experiment in the location of ethnic Germans. By law ethnic Germans (foreigners but with German ethnicity) are considered Germans. After the Iron Curtain fell, Germany saw a surge in immigration of ethnic Germans. Immigration authorities decided to allocated randomly these ethnic Germans into different counties for the period 1996-2001. Glitz (2006) finds that ethnic German immigration has no effect on wages and a negative effect on employment, although this effect disappears once controlling for selection into labor markets is included. In sum, similarly to research in the U.S., immigration in Germany does not seem to have a drastic negative effect on natives' labor market outcomes. Rather, the effect on wages and employment is null.

The results of research done in the U.S. and Germany are encouraging in order to understand why immigration seems to have a close to zero effect on natives' labor market outcomes. My research project is at the heart of that puzzle. The mechanism of the impact of immigration is missing. Micro level data is needed to understand what immigrants do, and more fundamentally, the accounting of the benefits at the firm level and the type of tasks immigrants do in their jobs. Given the similitude between Germany and the U.S., the use of German data can shed light on the mechanism of the impact of immigration.

This paper is in a preliminary stage. There are several issues that I still need to address. Labor inputs could be endogenous to the production technology. A positive demand

shock causes a firm to hire more immigrants causing in an overstatement of the true effect of immigrants in the production process. Instrumental variables need to be used to correct for this problem. Another potential bias in my estimation is that my preliminary results are obtained from variation across plants within years and industries. The use of plant fixed effects will help to disentangle the effect of permanent effects of firms. For example, if low productivity firms hire more immigrants, then immigrant labor will be understated only because immigrants are in low productivity firms. In order to solve for this problem, we need to include a separate identifier for each plant such that the effect of immigration on productivity comes from variation within each plant. Another aspect that needs to be included is the restriction of the sample to manufacturing firms. Finally I need to address how the results help us think in the aggregate consequences of immigration. For example, the finding of immigrants being more productive than natives has important aggregate consequences. Using the Borjas (1995) traditional model, the gains of immigration rely on how wages respond to immigration. But if immigrants are more productive than natives, firms are "double-dipping" in the sense that not only they hire immigrants with a lower wage, but those that hire produce more than similar natives causing an upward shift in the labor demand. Hence, the benefits of immigration depend on the productivity of immigrants. Also, the mechanism of why and how firms take advantage of immigrant labor is still a black box and requires further research.

2 Model

The main goal of the paper is to estimate the marginal productivity of immigrants relative to natives. Marginal productivities are estimated from a production function of general form:

$$Y_{pt} = A_{pt} G\left[K_{pt}, QL_{pt}\right] \tag{1}$$

where Y and K represent sales and capital respectively and p refers to plant and t to year. QL represents the quality of labor variable. From now on, I will assume all labor inputs are perfect substitutes for each other. For example, Females, Immigrants and High Skilled workers are substitutes. In order to understand how the model works, first I assume that all labor inputs are equally productive. For simplicity, assume we can differentiate the workforce in terms of gender and nationality. Hence, QL_{pt} is defined as:

$$QL_{pt} = MN_{pt} + FN_{pt} + MI_{pt} + FI_{pt}$$

where M and F refers to male and female and N and I refer to native and immigrant. In this case the quality of labor restricts to the total number of workers in the plant $QL_{pt} = L_{pt}$ and the estimation only takes into account total labor and not the quality of labor. Instead of assuming all labor inputs are equally productive, assume labor inputs have different productivity:

$$QL_{pt} = MN_{pt} + \varphi_F FN_{pt} + \varphi_I MI_{pt} + \varphi_F \varphi_I \varphi_{FI} FI_{pt} \tag{2}$$

where φ_F, φ_I, and $\varphi_F \varphi_I \varphi_{FI}$ are the marginal productivities of females, immigrants and females immigrants relative to male natives. The literature on immigration has assumed all these terms are equal to one, but this need not be the case. In what follows I explain how to test for different marginal productivities. In order to simplify the estimation, as Hellerstein et al. (1999) does, I restrict equation (2) in two ways. First, I assume an equiproportionate restriction among two different inputs. For example, the proportion of female natives among females is equal to the proportion of natives in that plant ($FN/F = N/L$). The second restriction is that the ratio of marginal productivity of two inputs within one demographic group (i.e. females) is equal to the ratio of marginal productivity of the same two inputs within other demographic group. In other words, the marginal productivity of immigrants relative to natives among females is $\varphi_I \varphi_{FI}$ and as the marginal productivity of immigrants relative to natives among males is φ_I, the condition requires $\varphi_{FI} = 1$.

Imposing these two conditions we can define QL_{pt} as

$$QL_{pt} = \left\{ \left(L_{pt} + \left[\varphi_F - 1 \right] F_{pt} \right) \left(1 + \left[\varphi_I - 1 \right] \frac{I_{pt}}{L_{pt}} \right) \right\} \tag{3}$$

where F refers the number of females in the plant and I to the number of immigrants.[3] The key parameters are φ_F and φ_I. They refer to the relative marginal productivity of females and immigrants. Notice that if females and immigrants are equally productive to males and natives ($\varphi_F = \varphi_I = 1$) the term QL limits to L. Remember that females and immigrants are assumed to be perfect substitutes in the production function (1). In future versions of the paper, I will relax this assumption. The goal is to estimate φ_F and φ_I using this framework. In the empirical application, I include not only females and immigrants but also low and high skilled workers. The share of low and high skilled workers can affect the productivity of the plant in ways related to the share of females or immigrants. Future versions of the paper

[3] After doing some algebra $QL = (L + (\varphi_F - 1)FN + (\varphi_I - 1)MI + (\varphi_F \varphi_I - 1)FI)$ which can be expressed as $(L + (\varphi_F - 1)F(1 - I/L) + (\varphi_I - 1)I(1 - F/L) + (\varphi_F \varphi_I - 1)F * I/L)$. This can be rewritten as $(L + (\varphi_F - 1)F - (\varphi_F)F * I/L + (\varphi_I - 1)I - (\varphi_I)F * I/L + (\varphi_F \varphi_I + 1)F * I/L)$. Finally, this term is equal to $(L + (\varphi_F - 1)F + (\varphi_I - 1)I - F * I/L(\varphi_F - 1 + \varphi_I - \varphi_F \varphi_I)$. This expression leads to equation (3) in the text because $\varphi_F - 1 + \varphi_I - \varphi_F \varphi_I = (1 - \varphi_F)(1 - \varphi_I)$.

will include different specifications in the term QL_{pt}. Doing the same analysis as before, the inclusion of low and high skilled workers will modify the term QL_{pt} as follows:

$$QL_{pt} = \left\{ \left(L_{pt} + \left[\varphi_F - 1 \right] F_{pt} \right) \left(1 + \left[\varphi_I - 1 \right] \frac{I_{pt}}{L_{pt}} \right) \left(1 + \left[\varphi_B - 1 \right] \frac{B_{pt}}{L_{pt}} + \left[\varphi_W - 1 \right] \frac{W_{pt}}{L_{pt}} \right) \right\} \tag{4}$$

where B stands for blue collar workers and W for white collar workers. The omitted group is part-time workers and trainees.[4]

Using a Cobb-Douglas production function and taking logs to equation (1), we obtain

$$\ln Y_{pt} = \ln(\widetilde{A}_{pt}) + \beta_K \ln(K_{pt}) + \beta_{QL} \ln QL_{pt} + \varepsilon_{pt} \tag{5}$$

In this way the parameters φ_F and φ_I will give the marginal productivity of females and immigrants with respect to the omitted group (males and natives respectively). If marginal productivities are the same, we expect $\varphi = 1$. If immigrants are relatively more productive than natives, then $\varphi_I > 1$.

The main problem in estimating production functions is the endogeneity of inputs. It is reasonable to think that the firm takes an input decision when observing a productivity shock (Marschak and Andrews, 1944). A positive productivity shock causes an increase in the demand for labor, leading to believe that labor is too important in the production process. If this is the case, the estimates will be upward biased. On the other hand, suppose there are some firms that consistently hire more females or immigrants (say small and low wage firms). If this is the case, the estimates will be downward biased because unobserved components of sales will be negatively correlated to the share of females or immigrants. Hence, we will conclude spuriously that the marginal productivity of females or immigrants is too low just because they are segregated in low productivity firms.

The literature on the estimation of production function has tried to solve the endogeneity of inputs in different ways.[5] The first strategy is to use Instrumental variables that are correlated to inputs but not to unobserved components in the production function. A straightforward instrument is the use of input prices. However, there needs to be sufficient variation across plants in input prices in order for the prices to be a valid instrument. The second strategy relies in the use of plant fixed effects (Mundlak, 1961). The main assumption behind this procedure relies that unobserved productivity is not time invariant. In the

[4] Check the Appendix for variable definitions. I define Blue collar workers as those workers that are production workers and are dividided as Not Qualified, Craftsmen and Skilled; White Collar workers are defined as non-production workers (i.e Qualified workers).

[5] For a literature review on this topic see Ackerberg et al. (2005b).

example described above, suppose there are firms that hire females and immigrants just because they are low productivity. Including plant fixed effects solves this problem because the estimator will include only the labor input variation within each plant and will not consider plants are similar to each other. The drawbacks in including plant fixed effects are that inputs need to be strictly exogenous to obtain consistent estimates and also because fixed effects absorb important variation. The last procedure relies in a semiparametric approach first proposed by Olley and Pakes (1996). This procedure assumes labor is a variable static input and capital is a dynamic quasi-fixed factor. This means that labor is not endogenous only capital stock. They assume that labor is not correlated to previous decisions by the plant or unobserved shock components. Nevertheless, capital stock is correlated with unobserved components, but once investment in the previous period is taken into account, it is possible to estimate consistent estimates of labor and capital using a two stage procedure. Hence, assuming that unobserved productivity can be modelled as a semiparametric function of investment and capital, the coefficient on labor can be identified.[6]

It is important to mention that any method relies on different assumptions about the unobservable factors. Below, I will implement the Olley and Pakes (1996) procedure and in future work I will modify the procedure including instrumental variables for the labor inputs as suggested in Ackerberg et al. (2005a). The estimation of regression (5) is by Non-Linear Least Squares and in order to control for unobserved heterogeneity I use industry, region and year fixed effects.

The goal is to compare the marginal productivity estimates φ_F and φ_I to the relative wage of those groups. φ_F and φ_I provide only an estimate of the productivity of females and immigrants and if markets are competitive we expect this productivity to be equal to the wage paid to them. If we believe that natives and immigrants are perfect substitutes, the argument for hiring immigrants implies that immigrants put downward pressure on wages (or that they are consistently paid less than natives). Hence the appropriate test will be to estimate the following regression at the plant level:

$$\ln w_{pt} = \alpha + \ln QL_{pt} + \upsilon_{pt} \tag{6}$$

where $\ln w_{pt}$ is the log of total wages in plant p at time t and the quality of labor term is defined as

$$QL_{pt} = \left\{ \left(L_{pt} + \left[\lambda_F - 1 \right] F_{pt} \right) \left(1 + \left[\lambda_I - 1 \right] \frac{I_{pt}}{L_{pt}} \right) \right\} \tag{7}$$

[6] Olley and Pakes (1996) are interested in estimating the coefficients on labor and capital. As opposed to their paper, I am interested only in identifying the coefficient on labor. As such, I only estimate the first stage in their procedure such that I can recover the labor coefficient.

where the coefficient λ represents the relative wage of that group with respect to the omitted group. The term λ represents how females and immigrants are underpaid or overpaid with respect to males and natives respectively. $\lambda > 1$ implies the sociodemographic group is paid more than the omitted group and similarly for $\lambda < 1$. If factor markets are competitive, then $\lambda = \varphi$. Hellerstein et al. (1999) argue that $\lambda < \varphi$ is evidence in favor of discrimination in the labor market given that inputs are not paid their relative contribution to production in the plant. Instead of arguing in favor of discrimination, I just recognize a gap between productivity and wages. This could be driven by hiring costs for example. Nonetheless, the fact that productivity varies among different groups has important aggregate consequences.

Hellerstein et al. (1999) run regression (6) at the plant level. They do this mainly for two reasons: 1. The wage reported in the Census is not directly comparable to total wages for the plant, and 2. Their matched sample represents only 12% of the workforce. In contrast to their dataset, we have access to the full workforce and the wage reported is the one paid by the firm. I present results not only using wage aggregation at the plant level, but also I use individual data to obtain estimates of λ in order to test the robustness of the results. As λ is just the relative wage ($\lambda_F = \frac{w_F}{w_M}$ or $\lambda_I = \frac{w_I}{w_N}$), I estimate a regression using log wages at the individual level:

$$\ln w_{ipt} = \alpha + \beta_F F_{ipt} + \beta_I I_{ipt} + \varepsilon_{ipt} \tag{8}$$

where F and I are indicator variables and the constant represents the average wage of the excluded group in the plant (native males). In order to recover λ, a transformation $\lambda = \exp(\beta)$ is used. Regression (8) is simpler than regression (6) at the individual level:

$$\ln w_{ipt} = \alpha + \ln QL_{ipt} + \upsilon_{ipt} \tag{9}$$

because the equation is nonlinear in the parameters of interest ($QL_{ipt} = \{(1 + [\lambda_F - 1] F_{ipt}) (1 + [\lambda_I - 1] I_{ipt}$ As the dataset used in this study is large at the individual level, I will use regression (8) to estimate the parameters of interest.[7]

Regressions (6) and (8) have the same possible biases as the estimation of the production function (5), so the solutions to this problem are similar to that case. However, the goal of the paper is to estimate the productivity of immigrants and contrast it with the wage they are paid. If both regressions are biased, we expect the bias to be in the same direction.

[7] Future work will include plant fixed effects. As in the case of the estimation of production functions, plant fixed effects will help in identifying the coefficients because it is likely that there are low productivity firms that pay lower wages. Hence, not including fixed effects will bias the estimates just because immigrants and females are located consistently in specific type of firms. Now, another advantage of using regression (8) instead of regression (9) is that the inclusion of plant fixed effects slows the convergence process substantially.

In order to check the sensitivity of the results to these biases, in future work I will explore different specifications (mor labor inputs or more disaggregation) as well as to include plant fixed effects.

3 Data

I use the LIAB data from Germany. This is a matched employer-employee dataset that links information for the firms in the Establishment Panel Dataset (IAB) with workers in the Employment Statistics Register (Social Security Records) from 1993-2004.[8] The Establishment Panel Data (IAB) is an annual survey of German establishments, administered since 1993 by Infratest Burke Sozialforschung. The establishment panel is based on a stratified random sample with respect to 10 categories of the establishment size and 16 categories of the industry from the population of all establishments and only includes establishments with at least one employee covered by social security. In 1993 the sample included 4,265 plants accounting for 0.27% of all plants in West Germany and 11% of total employment. Since 1996 East Germany is included, and the sample size increased to 8,879 plants. The sample size has increased since then and in 2004 it covered 19,234 plants. Plants are kept in subsequent years only if they are still considered representative and if the plant has not closed. Some of the variables included in the panel data set are: number of employees, investment, sales, overall wage bill, technological status, assessment of overall company economic situation, establishment size and industry.

The IAB data is matched to information on individuals from the German Employment Register which contains information on all employees and trainees subject to social security taxes. By law employers have to provide information to the social security agencies for those employees registered by the social security system. Excluded from the sample are self employed, civil servants, family workers and students enrolled in higher education. Among the variables that employers are obliged to declare about their workers are occupation, gender, year of birth, nationality, marital status, number of children, and schooling. Other labor market variables include: start and end of each employe notification and average daily wage for an employment spell. As the employer is not required to ask for this information, we expect occupation and gender to be precisely measured but education to be measured with some error, especially for immigrants, given that in other countries have a different educational system than in Germany. For this reason, I proceed to analyze labor groups in terms of occupation and not in terms of education.

One drawback in the IAB data is that it does not have a capital stock measure. The

[8] See Alda et al. (2005) and Andrews et al. (2004) for more details about the LIAB dataset.

dataset only includes investment expenditures. Previous research has used the sum of current and previous investment as a proxy for capital stock.[9] Instead of following this approach, I construct a proxy for capital stock based on four investment periods and sales growth. This procedure is valid only for plants that are present at least four years in the sample. For the rest of the plants, I multiple impute capital stock for the initial period.[10] The Appendix contains full details in the procedure.

Another disadvantage of the dataset is the measure of sales. The dataset includes a variable that measures the percentage of intermediate costs, but around 50 percent of plants do not report this variable. Instead of including more noise to the data, I decide not to transform sales into a value added specification. Although previous literature has emphasized the benefits of such transformation, Basu and Fernald (1997) argue that the value added specification is valid only if we assume there is perfect competition, absent this aspect we could make things worse by including a value added specification.

Before the cleaning procedure, we have information on 138,431 year observations and around 24 million worker observations. The Appendix includes exact details about the cleaning procedure. I restrict the sample to those firms that declare at least 15 employees in the Social Security records in all years and I drop all firms in which the number of workers from the Social Security records differs by more than 30% from the IAB dataset. I drop those plants that do not declare sales as their turnover measure (mainly financial institutions) and industries like Recycling, Utilities, Public Administration, Finance, and Household Services. As the IAB changed their sampling procedures in 1996 (East Germany is included and more smaller establishments), I use data since 1996 to avoid problems of comparison between sampling procedures. My final sample consists in 25,454 plant-year observations and 5,236 different plants (22,858 with Imputation procedure), the average duration of a plant in the dataset is 3.9 years.

Table 1 shows some basic descriptive statistics using the sample weights. For simplicity, I just present statistics for three years: 1996, 1999, 2002. In general, all variables are fairly constant throughout the period of analysis. The number of workers is around 80 workers for the three years and their average age is close to forty years old. The proportion of workers is fairly constant among immigrants and females. Native females represent one-third of natives and female immigrants represent close to one-third of immigrants. However, immigrants are not equally represented in the occupational structure. Immigrants are predominantly low skilled.[11] While 90 percent of immigrants are in low skilled occupations, natives only

[9] Addison et al. (2005) uses the sum of current and lagged investment as a proxy for capital stock and Addison et al. (2003) uses replacement investment.

[10] I follow the procedure described in Rubin (1987) and Rubin and Little (2002).

[11] Only for this part I include trainees, part time and blue collar workers together as a single group. In

account for 35 percent of the same group. Nevertheless their disadvantage in the occupational structure, the wage gap between immigrants and natives is not large (around 3 percent). Among plant characteristics, the sample is fairly representative of four regions in Germany. Using the sample weights, small firms represent 65 percent of the total number of plants. Immigrants are not hired only by a few firms, around 65 percent of the plants hire immigrants and, among those with positive immigrant employment, the share of immigrants in the workforce is around 10 percent. Immigrants are not equally located through all Germany. Immigrants in West Germany represent between 8 and 10 percent of the workforce, while in East Germany they represent less than one percent of the workforce.

4 Preliminary Results

The final goal of this paper is to estimate the relative productivity of immigrants with respect to natives and contrast it with the relative wage of immigrants and natives. The regressions that follow are in preliminary stage. In the interpretation of these results we need to keep in mind the following:: I have not corrected standard errors for the use of a generated variable (Capital stock), the regressions do not include plant fixed effects, and finally, I have not experimented with different functional forms. Nonetheless, the preliminary estimates will give us an idea of how the estimators behave without too much structure in the model.

Table 2 presents the main results in the estimation including year, industry and region fixed effects.[12] Using regression (5) and quality of labor term (4), the first three columns (1)-(3) show the relative marginal productivity of immigrants, females, blue and white collar workers. Column (1) includes all four regions while Column (2) only includes West Germany. Immigrants are relatively more productive than natives in Germany as a whole but when we restrict the regression to only West Germany immigrants are not relatively more productive than natives.[13] Immigrants are 20 percent more productive than natives in Germany but only 5 percent more productive in West Germany and this result is not statistically significant from one. One of the reasons why the result is not robust for West Germany could be that plants located in this area are more productive than East German plants. As we mentioned earlier, if there are some type of plants that consistently rely more on immigrants, the estimates will be biased, and as such the use of plant fixed effects will eliminate that type of bias.

On the other hand, females are consistently less productive than males. Females are

the analysis below, I refer as low skilled workers only to blue collar workers.

[12] Remember that the standard errors have not been corrected by the fact that capital stock is a generated regressor from the Multiple Imputation procedure.

[13] Remember that $\varphi > 1$ implies the group is relatively more productive than the excluded group.

around 55 percent less productive than males. This result is fairly constant even if we restrict the sample to only West German plants. In results not shown in Table 2, I relax the assumption of equal marginal productivity between females natives and immigrants. The results are robust to this transformation. Column (3) presents the results using the Olley and Pakes (1996) procedure for West Germany (only correcting for the endogeneity of capital stock) restricting the sample only for those plants with positive investment. The coefficients do not vary too much and the same conclusion arises for females and immigrants. The finding of females being less productive than males confirms the results in Hellerstein et al. (1999), they find that females are only 16 percent less productive than males for the U.S. The magnitude of the estimate is surprising. It is likely that low productivity firms employ more females. In fact, in my sample females are overrepresented in low wage firms. Around 60 percent of the workforce among the lowest wage firms employ women.[14]

The relative marginal productivities need to be contrasted to the relative wages. Columns (4)-(5) in Table 2 show the estimation using total wages in the plant instead of sales, in particular, it shows regression (6) using industry, region and year fixed effects. Column (4) includes all Germany and Column (5) includes only West Germany. The coefficients are more precisely estimated than in the case of the production function. Immigrants are paid 10 percent more than natives in all Germany, but underpaid 2 percent in West Germany (although the coefficient is not statistically different from one). This difference could be driven by systematic differences between the two regions. As wages are higher in West Germany than in East Germany and as immigrants are not present in the Eastern economy, immigrants look like they earn higher wages than in natives. The use of plant fixed effects would help in disentangle the permanent effect of a permanent firm characteristic. This paper is interested in contrasting relative productivity with relative wages. The hypothesis test for this case $\lambda = \varphi$ is not rejected. We cannot reject the null hypothesis of immigrants being similar to natives in terms of productivity and sales.

As opposed to results in the U.S., females are paid more than their marginal product. Females are paid only 33 percent less than males. This result is statiscally significant different from one. Hellerstein et al. (1999) find that females are underpaid 45 percent with respect to males. In Germany, females are paid 20 percentage points more than their marginal product while in the U.S. they are paid 30 percentage points less. This difference is striking and deserves further attention. This result could have important consequences for the wage structure of Germany and U.S. If women are paid more than their marginal product, the

[14] I do not present a table for this result. I obtained the median wage paid by each firm, then I sort the firms according to this median wage and assign them into quintiles. The workforce of the first quintile is around 60 percent female. In contrast, the workforce of the fifth quintile is around 19 percent female. The same is not true for immigrants. The share of immigrants is fairly constant accross the quintiles.

female wage gap is potentially less than in the case of a wage similar to the marginal product.

The next set of results concern blue and white collar workers. Blue collar workers are between 50 and 100 percent more productivities than trainees and part time workers. The wage results show that they are paid their marginal product. The results are almost identical. If immigrants and blue collar workers are paid their marginal product and females are paid more, how do firms make profits? The last row in table 2 shows the answer. White collar workers (qualified workers) are more than 400 percent productive than trainees and part time workers. However, they are paid only twice as much, and this difference is statistically significant. As in the case of females, these results have important consequences for the wage structure in Germany. These results could potentially explain the evolution of wage inequality of Germany and the U.S. It is known that wage inequality has been increasing in the U.S. in the last twenty years, while in Germany this surge has not been as pronounced as in the U.S. Firms wage pay policies deserve a more thorough examination to explain their role in the surge of wage inequality in the U.S.

Columns (6) and (7) show the results for regression (8) at the individual level controlling for industry, region and year fixed effects. The number of observations is more than 4 million in both cases and the estimates are very precisely estimated. Immigrants are paid 3 percent less than natives and this is consistent for West and East Germany. Females are paid only 18 percent less than natives. But remember that females are more segregated in low wage firms than males, so it is likely that this effect is affecting the estimates in columns (4) and (5). Nevertheless, the result that females are paid more than their marginal product still holds. Also, white collar workers are still paid less than their marginal product. Future research needs to take into consideration the inclusion of plant fixed effects to solve for the problem of permanent segregation of females and immigrants into some type of firms.

5 Aggregate Consequences of Immigration

In one important paper in the immigration literature, Borjas (1995) mentions: "Ironically, even though the debate over immigration policy views the possibility that immigrants lower the wage of native workers as a harmful consequence of immigration, the economic benefits from immigration arise only when immigrants *do* lower the wage of native workers." (emphasis in original, p.10-11). This argument is valid if immigrants are equally productive to natives. If immigrants have higher productivity, immigration could increase the benefits to the host economy without relying on the assumption of lowering the wage of native workers.

In the future version of the paper, this section will include the aggregate consequences of the results presented in the previous section. I will try to include the same analysis as in

Borjas (1995) but including the results of relative marginal productivity and relative wage. However, the results shown in Section (4) depict a simple story. Firms that hire immigrants do not seem to pay lower wages to immigrants. Neither there is evidence on immigrants being more productive than natives. In sum, immigrants are similar to natives in terms of wages and productivity. Hence, if economic benefits from immigration are different from zero, they have to appear in non-wage compesation schemes or in the immigrant labor flexibility (no holidays; first hired, first fired, etc.).

6 Conclusions

My preliminary results imply that immigrants are as productive as natives and they are not systematically underpaid relative to natives. On the other hand, females are 50 percent less productive than males but are only paid 40 percent less. As oppose to the U.S. case, females are paid more than their marginal productivity. White collar workers are not paid their marginal product but less. All these results have important consequences for the wage structure of Germany. Moreover, the results shed light on why inequality has surged in the U.S. while in Germany has not increased by as much. Firm wage policies deserve further examination in their role to explain wage inequality.

If natives are similar to immigrants in terms of wages and productivity, the economic benefits of immigration, as defined by Borjas (1995), are close to zero. However, other channels need to be explored. If firms hire immigrants, a reason should exist on why they do so. Among the possible reasons are non-wage benefits (holidays), labor flexibility, and management practices (monitoring costs).

This paper starts with the assumption that immigrants are substitutes for natives independent on the qualifications of each worker. An easy way to relax this assumption is to estimate a Cobb-Douglas production function but with two labor inputs. Within each labor input it is possible to include a quality of labor term for immigrants and natives. In this sense, it is possible that immigrants are not similar to natives. A different question worth examining is the estimation of the elasticity of substitution between immigrants and natives instead of assuming this number. I am currently working in this project, and so far my estimation results show that immigrants and natives are imperfect substitutes. Needless to say, these results are in preliminary stage and a more detailed analysis is needed.

In my future research agenda I plan to implement different procedures. I will include plant fixed effects in the estimation. This process is complicated because I am estimating a nonlinear model. Another aspect that needs to be considered is the estimation of wages at the individual level and not at the plant level. Also, I need to correct the standard

errors of the coefficients because I am using a generated regressor (capital stock). More importantly, I need to include a section that explains the aggregate consequences of my results. Immigration could have important aggregate consequences if immigrants are more productive than natives or they are underpaid, but so far I have not found that. I need to look at other channels as to why immigrants are hired. If immigrants are not paid less and do not produce more, it is hard to make a case as to why immigrants are hired in the first place. The simplest answer to this puzzle is that they are available, but this hypothesis implies that job matching costs for the firm are substantial for them not to hire a native worker.

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer, "Structural Identification of Production Functions," Working Paper 2005.

_ **, Lanier Benkard, Steven Berry, and Ariel Pakes**, "Econometric Tools for Analyzing Market Outcomes," Working Paper 2005.

Addison, John T., Lutz Bellmann, Thorsten Schank, and Paulino Teixeira, "The Demand for Labor: An Analysis Using Matched Employer-Employee Data from the German LIAB. Will the High Unskilled Worker Own-Wage Elasticity Please," IZA Discussion Papers 1780, Institute for the Study of Labor (IZA) September 2005. Available at http://ideas.repec.org/p/iza/izadps/dp1780.html.

_ **, Thorsten Schank, Claus Schnabel, and Joaquim Wagner**, "German Work Councils in the Production Process," IZA Discussion Papers 812, Institute for the Study of Labor (IZA) June 2003. Available at http://ideas.repec.org/p/iza/izadps/dp812.html.

Alda, Holger, Stefan Bender, and Hermann Gartner, "The Linked Employer-Employee dataset of the IAB (LIAB)," IAB Discussion Papers 6/2005, Institut für Arbeitsmarkt-und Berufsforschung 2005.

Andrews, Martyn, Thorsten Schank, and Richard Upward, "Practical Estimation Methods for Linked Employer-Employee Data," IAB Discussion Papers 3/2004, Institut für Arbeitsmarkt-und Berufsforschung 2004.

Basu, Susanto and John G. Fernald, "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, April 1997, *105* (21), 249–283.

Bonin, Holger, "Wage and Employment Effects of Immigration to Germany: Evidence from a Skill Group Approach," IZA Discussion Papers 1875, Institute for the Study of Labor (IZA) December 2005.

Borjas, George J., "The Economic Benefits from Immigration," *The Journal of Economic Perspectives*, 1995, *9* (2), 3–22.

_ , "The Economic Analysis of Immigration," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, May 1999, chapter 23, pp. 1697–1760.

_ , "The Labor Demand Curve Is Downward Sloping: Reexamining The Impact Of Immigration On The Labor Market," *The Quarterly Journal of Economics*, November 2003, *118* (4), 1335–1374.

Butcher, Kristin F., "Black Immigrants in the United States: A Comparison with Native Blacks and other Immigrants," *Industrial and Labor Relations Review*, January 1994, *47* (2), 265–284.

Card, David, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, January 2001, *19* (1), 22–64.

Chin, Rita, *The Guest Worker Question in Postwar Germany*, Cambridge University Press, 2007.

Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter, "Imputation Rules to Improve the Education variable in the IAB Employment Subsample," Discussion Paper 05-10, Center for European Economic Research 2005. Available at ftp://ftp.zew.de/pub/zew-docs/dp/dp0510.pdf.

Glitz, Albrecht, "The Labor Market Impact of Immigration: A Quasi-Experimental Evidence," Working Paper, UCL December 2006.

Göktürk, Deniz, David Gramling, and Anton Kaes, eds, *Germany in Transit. Nation and Migration 1955-2005* University of California Press 2006.

Hellerstein, Judith K., David Neumark, and Kenneth R. Troske, "Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations," *Journal of Labor Economics*, July 1999, *17* (3), 409–446.

Herbert, Ulrich, *A History of Foreign Labor in Germany, 1880-1980*, The University of Michigan Press, 1990.

Letterie, Wilko A. and Gerard A. Pfann, "Structural Identification of High and Low Investment Regimes," *Journal of Monetary Economics*, 2007, p. 1. Forthcoming.

Marschak, Jacob and William H. Andrews, "Random Simultaneous Equations and the Theory of Production," *Econometrica*, July 1944, *12* (3/4), 143–205.

Mundlak, Yair, "Empirical Production Functions Free of Management Bias," *Journal of Farm Economics*, February 1961, *43* (1), 44–56.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, November 1996, *64* (6), 1263–1297.

Ottaviano, Gianmarco I.P. and Giovanni Peri, "Rethinking the Effects of Immigration on Wages," NBER Working Papers 12497, National Bureau of Economic Research August 2006.

Pischke, Jörn-Steffen and Johannes Velling, "Employment Effects of Immigration to Germany: An Analysis Based on Local Labor Markets," *The Review of Economics and Statistics*, November 1997, *79* (4), 594–604.

Rubin, Donald B., *Multiple Imputation for Nonresponse in Surveys*, John Wiley and Sons, 1987.

_ **and Roderick J.A. Little**, *Statistical Analysis with Missing Data*, second ed., John Wiley and Sons, 2002.

Verick, Sher, Wilko A. Letterie, and Gerard A. Pfann, "Non-Linearities in the Expansion of Capital Stock," IZA Discussion Papers 1132, Institute for the Study of Labor (IZA) May 2004. Available at http://ideas.repec.org/p/iza/izadps/dp1132.html.

Table 1. Sample Description

Variables	1996	1999	2002	Variables	1996	1999	2002
log Sales	15.5	15.5	15.5	Num. Workers	76.72	78.21	82.34
sd	[1.26]	[1.26]	[1.29]	% Female	0.31	0.32	0.33
log K	13.8	14	14.1	% Immigrant	0.07	0.07	0.07
sd	[1.63]	[1.62]	[1.64]	Age	39.5	40	40.7
Firm Size				%Male-Native	0.64	0.63	0.62
15-50	0.65	0.66	0.64	%Female-Native	0.29	0.3	0.31
51-100	0.19	0.19	0.2	% Female-Immig	0.017	0.016	0.019
101-200	0.09	0.09	0.1	%Male-Immig	0.057	0.052	0.053
+200	0.07	0.06	0.06	%Native-Low	0.60	0.62	0.61
Region				%Native-High	0.32	0.31	0.32
North	0.15	0.14	0.15	%Immig-Low	0.065	0.06	0.06
Center	0.36	0.32	0.33	%Immig-High	0.008	0.008	0.012
South	0.25	0.31	0.29	Hire Immigrant	0.65	0.61	0.63
East	0.24	0.26	0.23	Wage Native	92.7	91.7	93.5
N	1,968	2,312	3,463	Wage Immig	90.3	91.4	91.3

Note: Calculations by the author.

Table 2. Regression Coefficients

	Production				Wages			
	(1)	(2)	(3)		(4)	(5)	(6)	(7)
	All	WEST	OP		All	WEST	All	WEST
φ_I	1.205	1.056	1.033	$\boldsymbol{\lambda}_I$	1.099	0.976	0.975	0.963
	[0.067]	[0.069]	[0.090]		[0.015]	[0.013]	[0.0005]	[0.0005]
φ_F	0.448	0.432	0.452	$\boldsymbol{\lambda}_F$	0.668	0.661	0.837	0.824
	[0.015]	[0.021]	[0.028]		[0.005]	[0.006]	[0.0003]	[0.0003]
φ_B	1.532	1.783	2.361	$\boldsymbol{\lambda}_B$	1.598	1.683	2.268	2.243
	[0.096]	[0.206]	[0.426]		[0.020]	[0.029]	[0.001]	[0.001]
φ_W	5.215	6.180	8.633	$\boldsymbol{\lambda}_W$	2.833	2.823	3.117	3.043
	[0.303]	[0.674]	[1.502]		[0.034]	[0.047]	[0.001]	[0.001]
N	22844	12767	9740	**N**	22844	12767	6474826	4941162
Inv>0			Y					
West		Y	Y			Y		Y
Individuals							Y	Y

Notes: Calculations by the author. Standard errors in parenthesis. Inv>0 includes only observations with positive investment. Columns (6) and (7) are regressions at the individual level.

A Capital

I follow the method proposed by Letterie and Pfann (2007) and Verick et al. (2004) to obtain a measure of capital stock. From the IAB data, I keep those firms that are present at least four years and have valid information in sales and investment. I proceed as follows:

1. (a) Using the Perpetual inventory method $K_t = (1-\delta)K_{t-1}+I_{t-1}$, estimate recursively Capital stock as $K_T = \sum_{t=1}^{T}(1-\delta)^{t-1}I_{T-t} + (1-\delta)^T K_0$ where $T = 4$ in our case.

   (b) As Letterie and Pfann (2007), assume capital stock grows at the same rate as sales: $K_T = \prod_{t=1}^{T}(1+g_t)K_0$. Here I smooth sales by the average in the four years and calculate $K_T = (1+\overline{g})^T K_0$

   (c) Solving the two equations and two unknowns we get $K_T = \frac{\sum_{t=1}^{T}(1-\delta)^{t-1}I_{T-t}}{1-\frac{(1-\delta)^T}{(1+\overline{g})^T}}$ and $K_0 = \frac{\sum_{t=1}^{T}(1-\delta)^{t-1}I_{T-t}}{(1+\overline{g})^T-(1-\delta)^T}$. In order to guarantee positive capital stocks, I restrict $\overline{g}$ to be non-negative.

   (d) In order to keep as many observations as possible, I keep K_0 and take it as a good estimate of capital stock in the initial period. I use the perpetual inventory method to calculate $K_1, ..., K_t$.

   (e) Firms with less than four years have missing values in capital. In order to maximize the sample size, I multiple impute capital stock for those plants using labor, sales, industry and state as explanatory variables under the Missing at random assumption. Rubin (1987) and Rubin and Little (2002) propose that the multiple imputation model has to be richer than the analyst model, and argue that not including the dependent variable (sales) can seriously biased the estimated of capital on sales. I multiple impute capital only for the first period and then I use the perpetual inventory method to calculate following capital stocks.

B Variables

All variables are in 2004 Euros.

- Investment: Investment is deflated by the Investment Price Index given by the Statistical Office (Statistiches Bundesamt) in its series Fachserie 17, R2, 1/2007. I use Euros instead of Deutsche Mark.

- Sales: Sales are deflated using a two digit NACE industry classification (28 industries). This classification starts since 1995. The Index is found in the Series: "Preise. Index der Erzeugerpreise gewerblicher Produkte (Inlandsabsatz) nach dem Güterverzeichnis für Produktionsstatistiken" published by the Statistical Office. For services, I use the Consumer Price Index.

- Labor: This variable is obtained from the Social Security Records.

  - Immigrant: All individuals with different nationality than German are considered Immigrants. In some cases, the plant declares that the worker is immigrant and the following year the same worker is native. I consider measurement error as those workers classified as Germans and in different periods the same worker is classified as foreigner, and hence I construct a foreigner variable that is constant across time. This measurement error is very small, accounting for around 1% of all immigrants.

  - Occupation: I have two different measures of occupation. A three index category of occupation, and a general classification that divides the workforce in six categories: Trainees, Part time workers, White Collar workers, Not Qualified, Skilled workers and Craftsmen. I consider blue collar workers as the last three occupations.

  - Education: It is well known that the education variable has serious problems of measurement error, see for example Fitzenberger et al. (2005). The problem relies in that plants, in general, do not ask employees their education or it is difficult to infer for foreigners, as such in some cases we observe that education can decrease for some workers. The problem is exacerbated for Immigrants. Fifteen percent of immigrants have missing values in education. As such I prefer to use the occupation classification.

- Wages: Wages are converted into euros and deflated using the Consumer Price Index. Wages refer to the average daily wage in the employment spell of each individual.

- Industry: The IAB includes an industry classification. It is worth mentioning that this industry classification is not consistent over time. In some cases, it is possible to assign new plants in some industries to the old classification. The Social Security records include industry as well. I use the industry classification (WZ73) given in the Social Security records mainly for two reasons: 1. It is more detailed, and 2. Even though the classification is not consistent over time (the classification changed in 1999 to WZ93), I was able to match the industries from one classification to other using the German Classification of Economic Activities, Edition 1993 (WZ93) published by the Statistical Office.

- State: I use the state classification provided in the Social Security records. There are sixteen states.