IZA/CEPR 11TH EUROPEAN SUMMER SYMPOSIUM IN LABOUR ECONOMICS

Supported and Hosted by the Institute for the Study of Labor (IZA)

Buch, Ammersee

17-19 September 2009

Machines and Machinists: The Effect of Imported Capital on Wage Inequality

Márton Csillag and *Miklós Koren

Machines and machinists: The effect of imported capital on wage inequality*

Márton Csillag and Miklós Koren†

April 2009

Abstract

This paper estimates the effect of capital imports on the wages of a large, representative sample of Hungarian machine operators. Using linked employer-employee data and detailed product- and firm-level import data, we match the precise occupation description of each worker to the equipment imported by their employers. We find that machine operators working on imported machines earn 16 percent more than those working on domestic machines. Estimating a structural assignment model of heterogenous workers and machines, we conclude that about one third of this wage gap is due to the higher returns to skill on imported machines, and two thirds are due to the higher skill of imported machine operators. Our structural estimates also suggest that imported machines contributed substantially to the increase in wage inequality in Hungary. Our results highlight a novel mechanism of skill-biased technical change.

1 Introduction

This paper investigates the role of capital imports in shaping wage inequality. Most developing countries lack significant production of machinery, and hence imports from R&D abundant countries provide the single most important source of machinery and technology embodied in them. (See Eaton and Kortum (2001) and Caselli and Wilson (2004).) Capital

*This research is carried out within the Center for Firms in the Global Economy (CEFIG). We thank Péter Harasztosi for excellent research assistance. Financial support from the FP7 program SCIFI-GLOW is gratefully acknowledged.

†*Csillag*: KU Leuven. E-mail: marton.csillag@econ.kuleuven.be, *Koren*: Central European University, IEHAS and CEPR. E-mail: korenm@ceu.hu.

imports are also the most important channel through which skilled-biased new technologies can be adopted in these countries. If capital goods complement skilled labor in production, the marginal product of skilled labor increases once imported machinery becomes available, and hence the skill premium rises.

We estimate the effect of imported machinery on the wages of a large, representative sample of Hungarian machine operators. Hungary provides fertile ground for studying the evolution of inequality, because earnings inequality has increased dramatically during the phases of transition and liberalization (Kezdi 2002). We have selected machine operators because they clearly complement machines in the production process. If the machine breaks down, it renders the marginal product of its operator zero. If the operator is negligent, it drastically reduces the productivity of the machine. Moreover, we hypothesize that the *quality* of the machine and the *skill level* of its operator are also complementary. An expensive, high-quality machine requires a skilled, well-trained operator, and someone with less than adequate training or who is less attentive may do more harm to a high-quality machine and its output than to a low-quality one. This implies that higher-quality machines raise the demand for skill; a form of skill-biased technical change.

As an example, consider the case of a Hungarian mining company, which purchased a kaolinite mining equipment from Britain, in 1982. The seller sent a team of workers to install the machine and to train the miners in using it. The British workers operated the machine in its first four months, demonstrating that it can achieve a productivity gain of 30 percent. After the seller's team left, problems with the machine became rampant. The Hungarian miners were not as attentive, took frequent and long breaks on the job, and the foreman who was most knowledgable about the machine was often absent. (One would be tempted to characterize it as a case of "socialist work ethic.") After only ten months of use, the machine broke down irreparably.

We use linked employer-employee data for the year 2002. Employees represent a 10 percent quasi-random sample of all machine operators (about 16,000 workers). For each employee, we have a number of personal characteristics (such as gender, age, schooling, occupation), and their employer's basic financial information and their detailed export and import transaction data from the Customs Statistics. We can thus distinguish machine operators who work at a firm using domestic machines from those who work at firm having recently purchased an imported machine. Moreover, we match the precise occupation descriptions of each operator to a set of machines that he or she works on. For example, "printing machine operators" are matched with "machinery for type-setting; printing type, plates," as well as with "printing machinery."

Our findings show that workers at firms that import their specific machinery earn 16 percent more than workers with no access to imported machinery of their specific occupation. About one third of this wage differential is due to omitted firm characteristics. Importing firms may be more productive, better managed, and may be able to attract a better workforce. When we contrast machinists (e.g., printing machine operators) working at firms that import *their specific* machine (e.g., typesetting machine) to those working at firms that import machines *unrelated to their occupation* (e.g., computers and trucks), we find a wage gap of about 10 percent. This is our preferred estimate of the effect of imported machinery on wages.

The difference in wages reflect differences in skill as well as differences in the *returns to skill*. Among workers working on domestic machines, the wage gap between those holding a secondary degree (12 years of education) and those with vocational schooling is about 12 percent. Among those working on imported machines, the return to a secondary degree is 19 percent. That is, imported machines increase the returns to skill substantially. However, much of the skills of machine operators are *unobservable* and are only partially explained by formal schooling. This is important, because imported machines are operated by more skilled workers than domestic ones, and hence our estimated wage differential is the combined effect of increased returns to skill and unobserved skill differences.[1]

To address the issue of unobserved skill differences, we build a structural matching model along the lines of Roy (1951), Jovanovic (1998), and Costinot and Vogel (2009). Workers are heterogenous in (unobserved) human capital, while machines are heterogenous in their quality. Imported machines are, on average, higher quality than domestic ones, but we allow for quality heterogeneity within both types.[2] There is a fixed supply of both workers and machines, and they are matched one to one in an efficient market. Machine quality and worker human capital are supermodular in the production function: the returns to skill are higher on higher-quality machines. This immediately gives rise to *positive assortative matching*: the most skilled worker operates the most productive machine, the second most skilled the second most productive, etc.

[1]See DiNardo and Pischke (1997) and Entorf, Gollac and Kramarz (1999) in the context of the wage effects of computers estimated by Krueger (1993).

[2]The heterogeneity within machine types is crucial to explain the less-than-perfect sorting of workers into machine types. Without this heterogeneity, the least skilled worker on an imported machine would still earn more than the most skilled worker on a domestic machine. This strict sorting is strongly rejected by the data. See also Armenter and Koren (2009) on the quantitative importance of heterogeneity in an export-selection model.

The equilibrium wage of a worker reflects not only their own skills and the productivity of the machine they work on, but is also influenced by pool of other workers and machines. The presence of higher-quality machines leads to a greater return to skill and a higher average wage. It also raises wage inequality, because not all workers can work at high-quality machines. The ones gaining the productivity boost from these machines are exactly the high-skilled workers who already enjoyed higher wages.

We then parametrize our model and estimate its parameters structurally. There are three key parameters: (i) the dispersion in human capital, (ii) the dispersion in machine quality, and (iii) the average quality premium of imported machines. The intuition for how we can identify these parameters is as follows. First, the average premium of imported machines can be identified using a differences-in-differences approach. The earnings differential of two workers with similar levels of human capital is greater if this pair of workers work on imported machines, because there the return to skill is higher. By contrasting the wage inequalities of matched pair of workers, we can identify the quality premium of imported machines. Second, the dispersion in machine quality weakens the sorting of workers onto imported and domestic machines. At one extreme, with very low dispersion, all low-wage workers will work on domestic machines, and all high-wage workers will work on imported machines. At the other extreme, as dispersion goes to infinity, the quality of a machine will be uncorrelated with whether the machine is imported, and workers will be randomly allocated across machine types. By estimating the strength of sorting, we can identify the quality dispersion parameter. Finally, overall wage inequality is the result of both machine quality dispersion and human capital dispersion. Having estimated machine quality dispersion beforehand, we can obtain human capital dispersion from the residual wage inequality.

Our structural estimates suggest that, on average, imported machines raise the returns to skill by 26 percent. This implies that of the overall 16 percent wage gap, about one third are explained by the increased returns to skill, and the remaining two thirds are explained by worker selection. Of the overall wage dispersion, around 60 percent is explained by the dispersion in human capital, while the remainder is due to the heterogeneity in machine quality.

To quantify the effect of imported machines on wage inequality, we conduct the following counterfactual exercise in our estimated structural model. We remove all imported machines from the economy and replace them with domestic machines, their quality drawn from the estimated distribution of domestic machine qualities. This results in a machine pool that is inferior (in the first-order stochastic dominance sense) to the actual machine pool. We match these pool of machines with the unchanged pool of workers. Our results are as follows. Due to the inferior machine pool, average wages decrease by about 4 percent. The 90/10 ratio of wages reduces from 2.59 to 2.48 percent.[3]

Our results can shed light on why wage and income inequality has increased tremendously in developing countries, and why these increases have mostly coincided with periods of trade liberalization (see Goldberg and Pavcnik (2007), for a survey). Most researchers point to skill-biased technical change as an explanation. This explanation faces two challenges, however. First, it is not clear why the timing of the increase in inequality is coincidental with liberalization. (In fact, standard Heckscher–Ohlin trade theories would predict a *decline* in inequality after opening up to trade.) Second, developing countries suffered a much more rapid increase in inequality than the U.S., the U.K., or other industrialized countries.

We believe that developing countries adopted most of their skill-biased technology by purchasing machinery from countries on the technology frontier. Alfaro and Hammel (2007) show that liberalization in emerging economies was followed by a surge in capital imports. This can explain both regularities. First, the liberalization of capital imports is a necessary condition for technology adoption. Second, because new technologies can be *adopted* at a much faster rate than they can be *invented*, their effects on inequality will also be much faster.

Other papers have also found evidence for complementarities between technology adoption and globalization. Verhoogen (2008) finds that Mexican exporters upgrade both their technology and the skill of their workforce after trade liberalization. Bustos (2007) obtains similar findings for Argentina. Both of these papers, however, link technology upgrading to the *export decision* of firms. Our model relates it to *imported* capital good. The key difference is that the use of imported machinery can affect a larger set of workers.

More broadly, our findings that machine quality and worker skills are complementary lend support to the view that complementarities are an important feature of the development process (see Kremer (1993), and Jones (2008)). If skilled workers are required to operate new, better technologies, then the lack of adequate education and training is a barrier to the spillover of technologies. Moreover, if labor market institutions do not facilitate the efficient matching of workers with machines, aggregate productivity will be substantially lower (see Bénabou (1996)). Both effects make it harder for poor countries to catch up to the productivity frontier, magnifying differences in income per capita.

[3]Other measures of dispersion show a similar decreasing pattern.

2 Data and Empirical Patterns

We use linked employer-employee data. Employee data come from the Hungarian Wage Survey (WS), which contains a 10 percent quasi-random sample of all employees, recording their earnings, 4-digit occupation, education, age and gender. We use the 2002 wave, and limit our sample to machine operating occupations, resulting in about 20,000 employees. In our benchmark specification, we limit the sample to firms that have at least 50 employees, because the sampling procedure of the WS is somewhat different for smaller firms. Results are virtually unchanged if we include all firms. We drop all employees earning 50,000 Ft per month (the level of the minimum wage in 2002) or less, and those earning more than 450,000 Ft. Again, results are not sensitive to these adjustments. Our final sample includes 16,326 employees.

On employers we use standard Earnings Statement measures (overall employment, total sales), as well as detailed trade flows from the Customs Statistics (CS). The CS contains the universe of trading firms, recording their exports and imports in 4-digit Harmonized System product breakdown for all years from 1992 to 2003. For each worker in the WS sample, we can precisely identify the international transactions of his/her employer. In particular, not only do we see whether the employer imported any machinery in the past, we also see the specific equipment goods that it imported.

We match the 4-digit occupation codes (FEOR) to the 4-digit product codes (HS) to identify machines and their operators. For example, FEOR code 8127 covers "Printing machine operators." This code is matched with "Machinery for type-setting; printing type, plates" (HS code 8442), as well as with "Printing machinery" (8443), but not with "Weaving machines" (8446). Note that this is a many-to-many match: the average occupation is associated with 4.5 machines, and the average machine is associated with 2.8 occupations. Table 6 in the Appendix provides several examples of these matches.

For each worker, we create two measures of access to imported machinery. The *generic* measure takes the value of 1 if the employer of the worker imported any piece of capital good between 1998 and 2002, and 0 otherwise. The *specific* measure takes the value of 1 if the employer imported any of the machines specific to the worker's occupation between 1998 and 2002. It is important to note that because the CS dataset contains the universe of imports, these measures only take a 0 value if the firm has not imported capital, or at least, not directly. (If some firms import capital indirectly, then our estimates of a wage difference can be thought of as a *lower bound*.) In what follows, we refer to workers at a firm importing their specific machinery as "working on imported machines," and all other workers as "working on domestic machines."

Table 1: Number of workers (top row) and average wages (bottom row) across importers and non-importers

		Imports specific machine: No	Imports specific machine: Yes	Total
Imports some machine	No	4,004		4,004
		11.24		11.24
	Yes	6,177	6,145	12,322
		11.41	11.53	11.47
Total		10,181	6,145	16,326
		11.34	11.53	11.41

Table 1 describes the joint distribution of the two capital import measures. Around three quarters of all machine operators work at a firm with some imported capital. Roughly one half of these work at a firm with imported machines *specific* to their occupation. Using the specific measure, around one third of machine operators seem to be working on imported machines.

The table also reports the average log earnings of workers in each machine category. Workers at a firm with some imported capital earn about 24 percent more than those at non-importers. Even among importers, wages are higher for those workers whose specific machinery is imported. This wage difference is about 12 percent.

We classify the education of workers into three categories: no vocational schooling, vocational schooling and high-school degree or higher. As Table 2 shows, 30 percent of machine operators have no vocational schooling, around 50 percent have finished vocational school, and around 20 percent have finished high school. The distribution is broadly similar across the two machine types, with the exception that the fraction of high-school educated workers is somewhat higher among those working on imported machines.

Table 2: Levels of education on each machine type

Level of education	Domestic machines	Imported machines	Total
No vocational schooling	0.300	0.305	0.302
Vocational schooling	0.523	0.488	0.510
High-school degree	0.177	0.207	0.188

3 Estimating the Wage Differential

We are interested in the wages of these machine operators. Our preferred wage measure is the monthly gross earning of the worker. This includes the regular earning for May 2002 and 1/12 of the bonuses of 2001. Because wages are not reported at an hourly rate, it is important to account for part-time work and overtime. In all our specifications below, we control for the number of regular monthly hours, and indicators for whether the employee works part time and whether they received any compensation for overtime. To make wages comparable across occupations, we also include 4-digit occupation fixed effects.

Table 3: The wage effects of imported machines

Dependent variable is log monthly earning	(1)	(2)	(3)
Uses imported machine	0.159	0.102	0.128
	(0.021)	(0.023)	(0.023)
Works at capital importer		0.111	
		(0.018)	
Vocational schooling			0.042
			(0.011)
High-school degree			0.120
			(0.014)
Uses * vocational schooling			0.031
			(0.018)
Uses * high-school degree			0.068
			(0.025)
Observations	16326	16326	16326
R-squared	0.386	0.394	0.401
Controls			
Occupation fixed effects	Yes	Yes	Yes
Part-time, overtime indicators	Yes	Yes	Yes

We first estimate the wage differential across workers on domestic and imported machines. Table 3 reports the results. In column 1, we only include the controls mentioned above. Workers using imported machines earn 15.9 percent more than those using domestic ones. This number is unchanged if we control for personal characteristics of the worker (gender, age, education).

In column 2, we control for whether the worker's employer imported *any* machinery. In other words, we contrast machinists (e.g., printing machine operators) working at firms that import *their specific* machine (e.g., typesetting machine) to those working at firms that import machines *unrelated to their occupation* (e.g., computers and trucks). We find that the use of imported machinery corresponds to 10.2 percent higher wages. Workers in occupations unrelated to the imported equipment also benefit; they enjoy 11.1 percent higher wages than those at firms not importing any piece of equipment. This may reflect measurement error stemming from our matching procedure (e.g., a printing machine operator may also be using a computer), or may capture firm-level omitted variables. Importing firms tend to be better in many dimensions than non-importers (Bernard, Jensen and Schott 2009, Altomonte and Békés 2008, Halpern, Koren and Ádám Szeidl 2009), and may hence employ better workers.

Column 3 looks at the return to education by machine type. We interact the indicators of educational categories with the imported machine indicator. The omitted category is no vocational schooling. The returns to education are higher among people using imported machines. The return to vocational schooling is 4.2 percent on a domestic machine and 7.3 percent on an imported one. The additional return to a high-school degree is 12.0 percent on a domestic machine and 18.8 percent on an imported one.

That is, imported machines increase the returns to skill substantially. However, much of the skills of machine operators are *unobservable* and are only partially explained by formal schooling. This is important, because imported machines are operated by more skilled workers than domestic ones, and hence our estimated wage differential is the combined effect of increased returns to skill and unobserved skill differences. To address the issue of unobserved skill differences, we build a structural matching model in the next section.

4 A Model of Worker Selection

4.1 Workers

There is a continuum of workers with unit mass. Workers are indexed by $\phi \in [0, 1]$. Worker ϕ is endowed with human capital $h(\phi)$. Without loss of generality, we index workers such that

h is increasing. Also assume that h is continuously differentiable and $h' > 0$ for almost all ω.[4] We can define the distribution of human capital across workers as $F(h) = \phi : h(\phi) = h$.

Workers supply their human capital inelastically in a perfectly competitive market. A worker with human capital h earns a wage rate of $w(h)$. The $w()$ function is an equilibrium outcome and is the key object of interest.

For ease of notation, we assume that human capital has a bounded support, $[\underline{h}, \bar{h}]$. None of our results hinges on this assumption. Each worker can choose to work in another sector (or home production), earning w_0 irrespective of their skill. Because there is an unbounded mass of potential new unskilled workers, the least skilled worker will earn wage w_0, $w(\underline{h}) = 0$. This assumption is needed to pin down the overall *level* of wages, and has no bearing on wage *inequality*.

4.2 Machines

There is a unit mass of machines, indexed by $\psi \in [0, 1]$. Machine ψ has quality $\theta(\psi)$. Similarly, to the distribution of human capital, we assume that θ is continuously differentiable and $\theta' > 0$ for almost all ψ. We can define the distribution of quality across machines as $G(\theta) = \psi : \theta(\psi) = \theta$.

Machines are in fixed supply, and are indivisible.[5] The owner of each machine maximizes output minus labor cost. They take the wage schedule $w(h)$ as given.

4.3 Technology

A production unit consists of one machine and one worker.[6] When a machine of quality θ is matched with a worker of skill h, the resulting output is

$$A(\theta, h).$$

Assumption 1. *The function A is increasing in both arguments, and is twice continuously differentiable.*

Assumption 2. *The function A is supermodular, that is*

$$A(\theta', h') + A(\theta, h) \geq A(\theta', h) + A(\theta, h'),$$

[4]This ensures that the distribution of human capital is continuous and there are no mass points.

[5]In future work, we wish to analyze the decision of firms to purchase a given type of machine.

[6]This is without loss of generality. If machines differ according to the number of operators they require, we can simply redefine the units.

for all $\theta' \geq \theta$ and $h' \geq h$. If the function is twice continuously differentiable, this is equivalent to $A_{\theta h} > 0$.

Each machine produces the same product, the price of which we normalize to one. It is straightforward to generalize our results for the case when machines produce differentiated products. This requires a stronger assumption of log supermodularity on A (see Costinot and Vogel (2009)).

Machine owners look for a worker with the optimal level of human capital.[7] More human capital yields higher output, but is also more costly. Machine owners solve the following optimization problem:

$$\max_h A(\theta, h) - w(h). \tag{1}$$

4.4 Equilibrium

Definition 1. *An equilibrium is a matching function $\Theta : \mathbb{R} \to \mathbb{R}$ that maps worker skill to machine quality, and a function $w : \mathbb{R} \to \mathbb{R}$ of wages, such that*

1. *each worker is employed,*
2. *each machine is employed,*
3. *h solves solves* (1) *for a machine of type $\Theta(h)$.*

Proposition 1. *The equilibrium is characterized by* positive assortative matching*, that is,*

$$\Theta(h) = \theta[F(h)]. \tag{2}$$

Wages are implicitly given by the following partial differential equation:

$$A_h[\Theta(h), h] = w'(h), \tag{3}$$

with the boundary condition that $w(\underline{h}) = 0$.

Because of positive assortative matching, each worker is matched with a machine with equal rank. The worker with the highest skill works on the best machine, the second most skilled on the second best, and so on. Note that matching is independent of the exact shape of A, as long as A is supermodular, the same assignment will occur.

[7] As matching is efficient, the solution is equivalent if workers look for and purchase machines to work on. In this case, wages can be obtained as the rents of human capital.

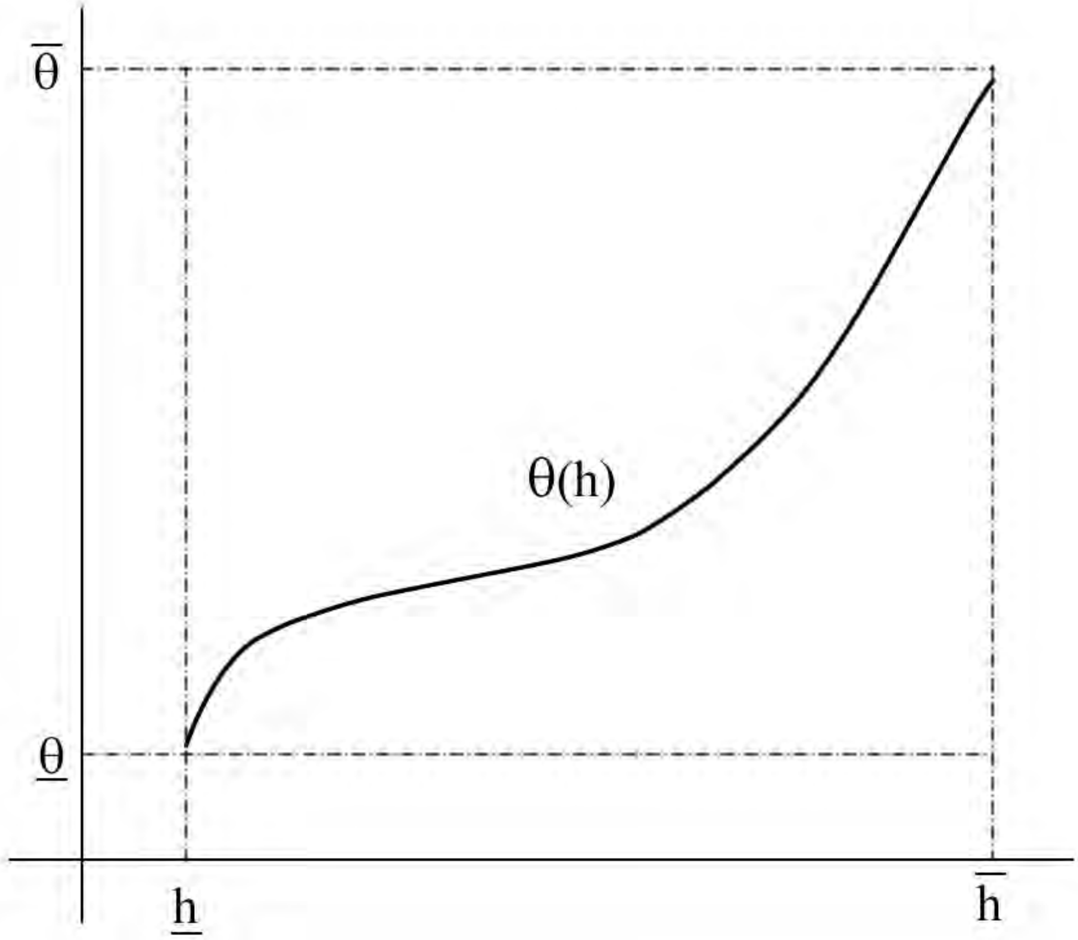

Figure 1: The matching function

Figure 1 displays a possible matching function $\Theta(h)$. It is continuously increasing in h, starts at $\Theta(\underline{h}) = \underline{\theta}$, and ends at $\Theta(\bar{h}) = \bar{\theta}$.

Wages depend on the marginal product of the worker on its current machine (which, in turn, depends on the matching function), but also on the skill and productivity of the rest of the workers. To see how, integrate (3) to obtain

$$w(h) = \int_{\underline{h}}^{h} A_h[\Theta(s), s]ds.$$

The wage of a worker with human capital h depends (through the matching function Θ) on the entire distribution of workers below this skill level, even conditional on $\theta(h)$, the machine that worker h is working on.

For several results, we also use the following assumption.

Assumption 3. *The function A is multiplicatively separable,*

$$A(\theta, h) = \theta h.$$

Note that this only assumes separability, because for any increasing functions f and g such that $A(\theta, h) = f(\theta)g(h)$, we could always redefine units of θ and h such that Assumption 3 holds.

If Assumption 3 holds, wages can be written as

$$w(h) = w_0 + \int_{\underline{h}}^{h} \Theta(s) ds. \tag{4}$$

Taken together, the wage equation and the matching function completely characterize the equilibrium.

4.5 Some Examples

To gain intuition about the equilibrium wage inequality, consider the following simple examples.

Two Machine Qualities. Suppose θ can only take two values: θ_H for a fraction α of the machines, and $\theta_L < \theta_H$ for a fraction $1-\alpha$. Positive assortative matching ensures that there is strict sorting of workers on the two types of machines: those with human capital between $\underline{h}$ and h^* work on low-quality machines, while those with human capital above h^* work on the high-quality one. The cutoff h^* can be determined implicitly from

$$F(h^*) = 1 - \alpha.$$

The assignment function is a simple step function,

$$\Theta(h) = \begin{cases} \theta_L & \text{if } h < h^*, \\ \theta_H & \text{otherwise.} \end{cases}$$

Integrating this with respect to h yields the wage function. Because the assignment function is a step function, the wage function is piecewise linear. We also know that it is continuous, which allows us to write

$$w(h) = w_0 + \begin{cases} \theta_L(h - \underline{h}) & \text{if } h < h^*, \\ \theta_L(h^* - \underline{h}) + \theta_H(h - h^*) & \text{otherwise.} \end{cases}$$

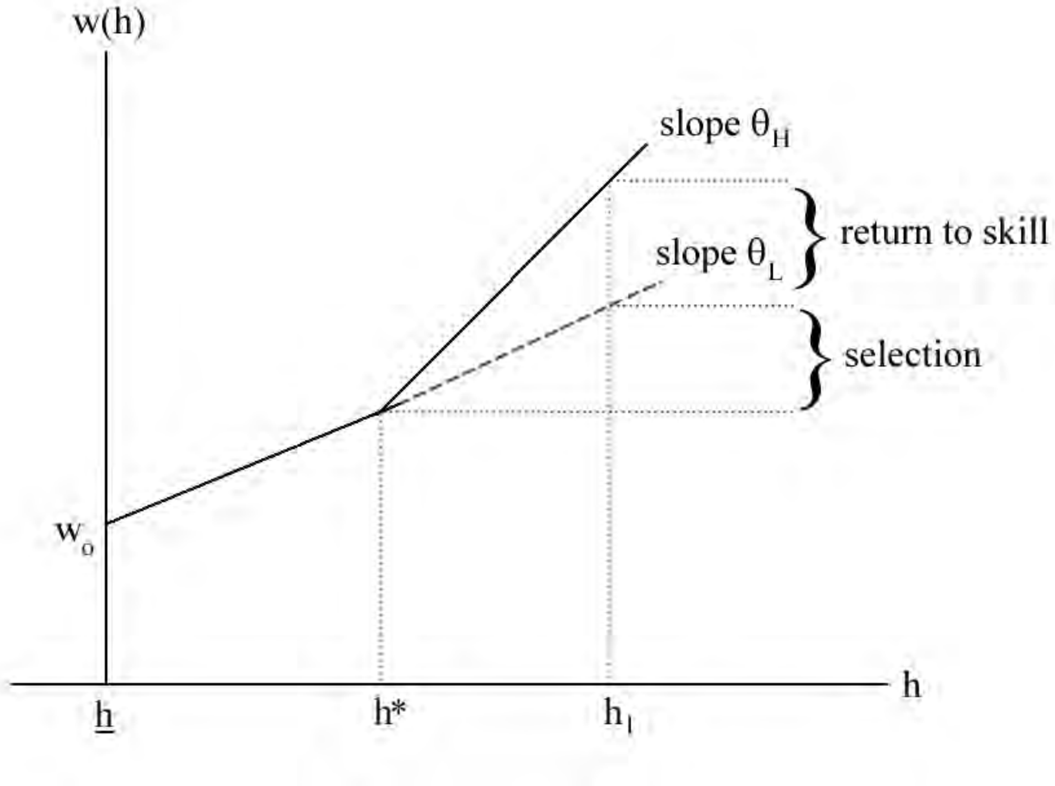

Figure 2: The wage function with two machine types

Since a fraction α of workers enjoy higher returns to schooling than the rest, both average wage and inequality increase in the share of high-quality machines, α. The average wage is

$$E(w) = w_0 + \theta_L[E(h) - \underline{h}] + \alpha(\theta_H - \theta_L)[E(h|h > h^*) - h^*].$$

Figure 2 displays the wage equation for this example. At $\underline{h}$, wages are zero. Below h^*, the slope of the wage function is θ_L, above h^* the slope is θ_H. Take a worker with human capital $h_1 > h^*$. They earn $\theta_H(h_1 - h^*)$ more than someone with human capital h^*. Part of this wage differential is due to the increased return to human capital, $(\theta_H - \theta_L)h_1$. The remaining $\theta_L(h_1 - h^*)$ is due to selection: this worker would earn more even if they worked on the low-quality machine. The goal of our structural model is to quantify the role of selection.

While this example is simple and intuitive, it has some strong counterfactual predictions. If there are only two types of machines, the sorting of workers by machine type should be perfect. Even the least skilled worker on a high-quality machine should earn more than the most skilled worker on a low-quality one. Empirically, this strict sorting property is clearly rejected. We therefore move on to discuss cases in which θ has a non-degenerate distribution.

Lognormal Machine Qualities. Suppose $G(\theta)$ is lognormal and $F(h)$ is normal such that

$$\ln\theta \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2),$$
$$h \sim \mathcal{N}(\mu_h, \sigma_h^2).$$

Inverting F, we get the human capital of individual ϕ as

$$h(\phi) = \mu_h + \sigma_h \Phi^{-1}(\phi),$$

where Φ is the cumulative distribution function of the standard normal distribution. Because of positive assortative matching we also have

$$\theta(\phi) = \exp[\mu_\theta + \sigma_\theta \Phi^{-1}(\phi)].$$

This allows us to write the matching function as

$$\Theta(h) = \exp\left[\mu_\theta + \frac{\sigma_\theta}{\sigma_h}(h - \mu_h)\right]. \tag{5}$$

We can integrate this to obtain the wage rate,

$$w(h) = w_0 + \frac{\sigma_h}{\sigma_\theta}\exp\left[\mu_\theta + \frac{\sigma_\theta}{\sigma_h}(h - \mu_h)\right],$$

or, in logs (assuming $w_0 = 0$),

$$\ln w(h) = \ln\sigma_h - \ln\sigma_\theta + \mu_\theta + \frac{\sigma_\theta}{\sigma_h}(h - \mu_h). \tag{6}$$

The return to human capital is

$$\frac{\partial \ln w}{\partial h} = \frac{\sigma_\theta}{\sigma_h}.$$

It is increasing in the dispersion of machine quality σ_θ^2, and decreasing in the dispersion of human capital σ_h^2. Recall that because of positive assortative matching, if worker A is one place ahead of worker B in the skill ranking, then A works on a machine one place ahead of that of worker B. The greater the dispersion in machine quality, the bigger the difference between the two machines, and the workers' marginal product.

The return to skill is decreasing in skill dispersion for the same reason. Worker A gets a better machine than worker B, irrespective of *how much* better he/she is. If skill dispersion is small, then these small differences get magnified by the large difference across machines.

Because h is normally distributed, wages are lognormal with log mean of $\ln\sigma_h - \ln\sigma_\theta + \mu_\theta$ and log variance σ_θ^2. The higher the dispersion of machine quality, the greater the wage inequality.

It may seem counterintuitive at first that skill heterogeneity, σ_h does not matter for wage inequality. There are two opposing forces at play. First, σ_h increases the dispersion in human capital and hence wages. Second, σ_h reduces the return to skill, because much better workers are now matched with only slightly better machines. With normal skill distribution, the two effects exactly cancel.[8]

5 Structural Estimation

To estimate the model, we need to relate machine quality θ and human capital h to observables. We take the approach that most variation in θ and h is unobservable, but certain categories of machines may differ systematically in their quality, and certain categories of workers may differ systematically in their skills. For example, while there is a large unobserved variation in machine quality, imported machines are, on average, higher quality than domestic ones. Similarly, workers with more years of schooling will possess more human capital on average.

More specifically, we divide machines into N distinct categories. In each category n, machines have a quality distribution G_n, so that the overall distribution of machine quality is

$$G(\theta) = \sum_{n=1}^{N} \alpha_n G_n(\theta),$$

where α_n is the share of machines falling in category n.

We similarly divide workers into K skill groups, with $F_k(h)$ denoting the distribution of human capital in skill group k. The aggregate distribution of human capital is

$$F(h) = \sum_{k=1}^{K} \beta_k F_k(h),$$

where β_k denotes the share of workers falling in skill group k.

For each individual ϕ, we observe their wage rate, $w(\phi)$, their skill group, $k(\phi)$, and the type of machine they are working on, $n(\phi)$. We only observe a cross section of workers.[9] Below we discuss the identification assumptions that allow us to estimate the model from these observables.

[8] As Heckman and Honore (1990) point out, this is a special case of the normal skill distribution, and does not generalize to log-convex distribution functions.

[9] We cannot link workers in different years in the Wage Survey, due to the lack of personal identifiers.

In our estimation, we use the conditional wage distributions within each machine type and skill group. These completely characterize all the observable characteristics of workers and machines.

5.1 Distributional Assumptions

Assumption 4. *There exists a G_0 distribution function such that*

$$G_n(\theta) = G_0(\theta/\mu_n).$$

This assumption ensures that the quality distribution of each machine type is of the same family, but with potentially different *scale* parameters. An example is a set of lognormal distributions with the same (log) variance but different (log) means.

Assumption 5. *There exists an F_0 distribution function such that*

$$F_k(h) = F_0(h - \lambda_k).$$

This assumption ensures that the human capital distribution of each skill group is of the same family, but with potentially different *location* parameters. An example is a set of normal distributions with the same variance but different means.

We also make use of the following first-order approximation:

$$w(h + \Delta) \approx w(h) + \theta(h)\Delta. \tag{7}$$

This follows from equation (4). The smaller the Δ, the more accurate the approximation.

Within each skill group k, we sort individuals by their wage rate. For each individual ϕ, we construct their rank within their skill group, denoting it by $p_k(\phi)$. Because wages are monotonically increasing in h, the ranking by wage is identical to the ranking by human capital. This is crucial, because we do not observe human capital directly.

We repeat this for machine types. The rank of an individual among workers working on the same type of machine is $p_n(\phi)$. Again, wages are increasing in machine quality θ (more precisely, wages are increasing in skill and more skilled workers are matched with higher quality machines), so this ranking is identical to the ranking by θ.

We can use the distribution assumptions 4 and 5 to express the unobserved human capital and machine quality of an individual as

$$h(\phi) = \lambda_{k(\phi)} + G_0^{-1}[p_{k(\phi)}(\phi)], \tag{8}$$

$$\theta(\phi) = \mu_{n(\phi)} F_0^{-1}[p_{n(\phi)}(\phi)]. \tag{9}$$

5.2 Estimating Average Machine Qualities

We estimate the average machine qualities from a difference-in-difference, or rather, a ratio-in-difference approach.

Take two skill groups, k and l. For each individual ϕ within skill group k, we find the individual ϕ' in skill group l who has the same rank,

$$p_k(\phi) = p_l(\phi').$$

According to (8), the difference in human capital between this pair of workers is

$$h(\phi') - h(\phi) = \lambda_l + G_0^{-1}[p_l(\phi')] - \lambda_k + G_0^{-1}[p_k(\phi)] = \lambda_l - \lambda_k.$$

We substitute this into equation (7),

$$w(\phi') - w(\phi) \approx \theta(\phi)(\lambda_l - \lambda_k). \tag{10}$$

where $\theta(\phi)$ is the quality of the machine that individual ϕ is working on. Let us denote this wage difference by $\Delta_k^l w(\phi)$. This wage difference can be calculated for each individual.

We then find an individual ϕ'' working on a different machine type m, who has the same rank as ϕ on machine type n,

$$p_n(\phi) = p_m(\phi'').$$

By Assumption 4, their respective machine qualities are a constant multiple of one another,

$$\frac{\theta(\phi'')}{\theta(\phi)} = \frac{\mu_m F_0^{-1}[p_m(\phi'')]}{\mu_n F_0^{-1}[p_n(\phi)]} = \frac{\mu_m}{\mu_n}.$$

Dividing $\Delta_k^l w(\phi'')$ by $\Delta_k^l w(\phi)$,

$$\frac{\Delta_k^l w(\phi'')}{\Delta_k^l w(\phi)} = \frac{\theta(\phi'')(\lambda_l - \lambda_k)}{\theta(\phi)(\lambda_l - \lambda_k)} = \frac{\mu_m}{\mu_n}.$$

We denote this ratio by $R_n^m \Delta_k^l w(\phi)$. We can calculate this ratio for each individual, and then estimate μ_m/μ_n as the mean (or median) across individuals. We also make the normalization that $\mu_1 = 1$.

5.3 Estimating G_0

The general idea behind estimating G_0 is that a more dispersed distribution of machine quality will lead to weaker sorting of workers. At one extreme, with very low dispersion, all low-wage workers will work on machines in the lowest quality group, and all high-wage workers will work on the highest quality machines. At the other extreme, as dispersion goes to infinity, the quality of a machine will be uncorrelated with its observed category, and workers will be randomly allocated across machine types. By estimating the strength of sorting, we can identify quality dispersion.

For a given G_0, once we estimated all the μ_ns, we can calculate all G_ns. This allows us to map a machine quality to each individual,

$$\theta(p_n) = \mu_n G_0^{-1}(p_n),$$

where, recall, p_n is the rank of the individual in the wage distribution of workers on machine type n.

The probability that a certain machine of quality θ belongs to machine type n is

$$\Pr(n|\theta) = \frac{\alpha_n g_n(\theta)}{g(\theta)}.$$

We work with the odds ratio of belonging to machine category n versus machine category m,

$$\frac{\Pr(n|\theta)}{\Pr(m|\theta)} = \frac{\alpha_n g_n(\theta)}{\alpha_m g_m(\theta)} = \frac{\alpha_n \mu_m g_0(\theta/\mu_n)}{\alpha_m \mu_n g_0(\theta/\mu_m)}.$$

For a given g_0, we can assign a odds ratio for each θ. We check if this odds ratio function equals the empirical odds ratio function, and if not, we pick another g_0.

In practice, we choose G_0 from a parametric family so that this functional fix point problem is easier to solve. However, we emphasize that our identification strategy can also be applied nonparametrically. More specifically, we let G_0 be lognormal with mean zero and variance σ_θ^2:

$$G_0(\theta) = \Phi(\ln\theta/\sigma_\theta),$$

where Φ is the cumulative distribution function of the standard normal distribution.

In this case, the log theta belonging to an individual with rank p_0 on machine type 0 is

$$\ln\theta(p_0) = \ln\mu_0 + \sigma_\theta\Phi^{-1}(p_0).$$

The log odds ratio can be written as

$$\ln\frac{\Pr(n|\theta)}{\Pr(m|\theta)} = C + \frac{\ln\mu_n - \ln\mu_m}{\sigma_\theta^2}\ln\theta = C' + \frac{\ln\mu_n - \ln\mu_m}{\sigma_\theta}\Phi^{-1}(p_0), \tag{11}$$

where C and C' are constants depending on the parameters α_n, α_m, μ_n, μ_m and σ_θ^2, but not on θ or p_0.

Equation (11) shows that machines with higher θ are more likely to belong to category n than m if and only if $\mu_n > \mu_m$. However, the speed with which this likelihood ratio is increasing in machine quality is decreasing in σ_θ. This suggests the following procedure:

1. For each individual, calculate p_0 as the rank their wage would have in the wage distribution of workers on machine type 0.

2. Estimate equation 11 by running an OLS of the log odds ratio on $\Phi^{-1}(p_0)$.

3. The coefficient is $(\ln \mu_n - \ln \mu_m)/\sigma_\theta$. Given our previous estimates of μ_n and μ_m, we can recover σ_θ.

5.4 Estimating Average Human Capital Across Skill Groups

Once we know $\theta(\phi)$, we can use equation (10) to estimate $\lambda_l - \lambda_k$ as

$$\frac{\Delta_k^l w(\phi)}{\theta(\phi)} = \lambda_l - \lambda_k.$$

We make the normalization that $\lambda_0 = 0$ (or any other constant).

5.5 Estimating F_0

The intuition behind estimating F_0 is that all wage dispersion not explained by θ should be due to differences in h.

According to equation (3), the derivative of wage with respect to human capital is

$$w'(h) = \theta(h).$$

The problem is that h is unobserved. However, we know the rank of each worker within their skill group,

$$p_k(h) = F_k(h).$$

This allows us to express the derivative of wage with respect to p_k as

$$\frac{dw}{dp_k} = w'(h)\frac{dh}{dp_k} = \frac{\theta(h)}{f_k[F_k^{-1}(p_k)]} = \frac{\theta(h)}{f_0[F_0^{-1}(p_k)]},$$

where the last equation follows from our Assumption 5.

Calculating the empirical wage gradient for each individual and dividing it by θ, we can obtain a nonparametric estimate of $f_0[F_0^{-1}(p_k)]$.

Again, we choose a convenient parametrization of F_0 as a normal distribution with mean zero and variance σ_h^2,

$$F_0(h) = \Phi(h/\sigma_h),$$

so that

$$\frac{dw}{d\Phi^{-1}(p_k)} = \sigma_h \theta(h). \tag{12}$$

The left-hand side of this equation is the wage gradient with respect to an inverse normal distribution function. We regress this on the previously estimated machine quality θ to obtain our estimate of σ_h.

The intuition is that the higher the σ_h, the dispersion of human capital within skill groups, the higher the wage difference between someone with a wage ranking of p_k and with $p_k + \Delta$. The wage difference also depends on machine quality θ, which is why we estimate σ_h in the last step.

5.6 Estimating w_0

So far we have only used information on wage *differences* across workers. All the above estimates are independent of the overall *level* of wages. We estimate w_0 such that the level of wages equals that in the data. More specifically, we integrate (12) for our estimated h and θ to obtain our estimate of wages, $\hat{w}$. We take the first percentile of $\hat{w}$ and subtract it from the first percentile of the actual w. The difference must come from w_0.[10]

5.7 Estimation Results

In our benchmark specification we implement the above procedure with two machine types, $N = 2$, and three skill groups, $K = 3$. Machines are split into domestic and imported.[11] Skill groups correspond to the level of former education of the worker. Group 1 contains people with no finished secondary education, group 2 contains workers with vocational schooling, and group 3 contains everybody with secondary or higher degrees.

Table 4 reports the estimated structural parameters. (Standard errors are bootstrapped.) Column (1) displays our preferred specification, where worker skill is divided into three subgroups (no vocational, vocational, secondary), and individual-level controls are occupation fixed effects.

All estimated parameters are plausible, and are highly statistically significant. The key parameter of interest is μ_2, the median (or log mean) machine quality. Recall that μ_1 has been normalized to 1. We estimate μ_2 to be 1.26, meaning that the median imported machine is 26 percent higher quality than the median domestic machine. We have chosen the units of

[10]Results are almost identical if we calculate this difference at other percentiles.

[11]Later we wish to further split imported machines by country of origin.

Table 4: Structural estimation results

Parameter	Description	(1)	(2)	(3)	(4)
mu2	Median productivity of imported machine	1.263	1.223	1.102	1.079
		(0.027)	(0.023)	(0.014)	(0.014)
sigma(theta)	Dispersion in machine quality	0.857	0.740	0.811	0.638
		(0.059)	(0.064)	(0.077)	(0.089)
lambda2	Mean human capital of workers with VS	0.038		0.041	
		(0.005)		(0.006)	
lambda3	Mean human capital of workers with SS or higher	0.115	0.092	0.117	0.093
		(0.008)	(0.008)	(0.007)	(0.007)
sigma(h)	Dispersion in human capital	0.299	0.291	0.305	0.295
		(0.018)	(0.003)	(0.007)	(0.004)
w0	Wage of lowest-skilled worker	0.532	0.520	0.536	0.520
		(0.009)	(0.008)	(0.011)	(0.013)
	Observations	16326	16326	16326	16326
Controls					
	Occupation fixed effects	Yes	Yes	Yes	Yes
	Part-time, overtime indicators	Yes	Yes	Yes	Yes
	Capital importer dummy	No	No	Yes	Yes

quality such that they correspond to the return to skill. In other words, the return to skill is 26 percent higher on imported machines than on domestic ones.

There is, however, considerable variation of machine quality within each machine type. The estimated $\sigma_\theta = 0.86$ implies that the interquartile range of domestic machine qualities is $[0.55, 1.82]$, while that of imported machines $[0.71, 2.36]$. Another way to quantify the magnitude of within-group variation is to conduct an analysis of variance. Of the total variation in $\ln\theta$, the import status of the machine explains only 2 percent.

Figure 3 displays the estimated density functions of machine quality for the two machine types. The solid line represents domestic machines. Because the quality of the median domestic machine has been normalized to 1, the distribution is centered around 0 on a log scale. Imported machines (dashed line) have on average higher quality, but there is a large overlap in the range of qualities across the two types of machines.

We estimate that vocational schooling adds 3.8 units of human capital, while degree granting secondary schooling adds another 7.7. The units of human capital are such that they bring an additional unit of wage on the median domestic machine. Because mean wages

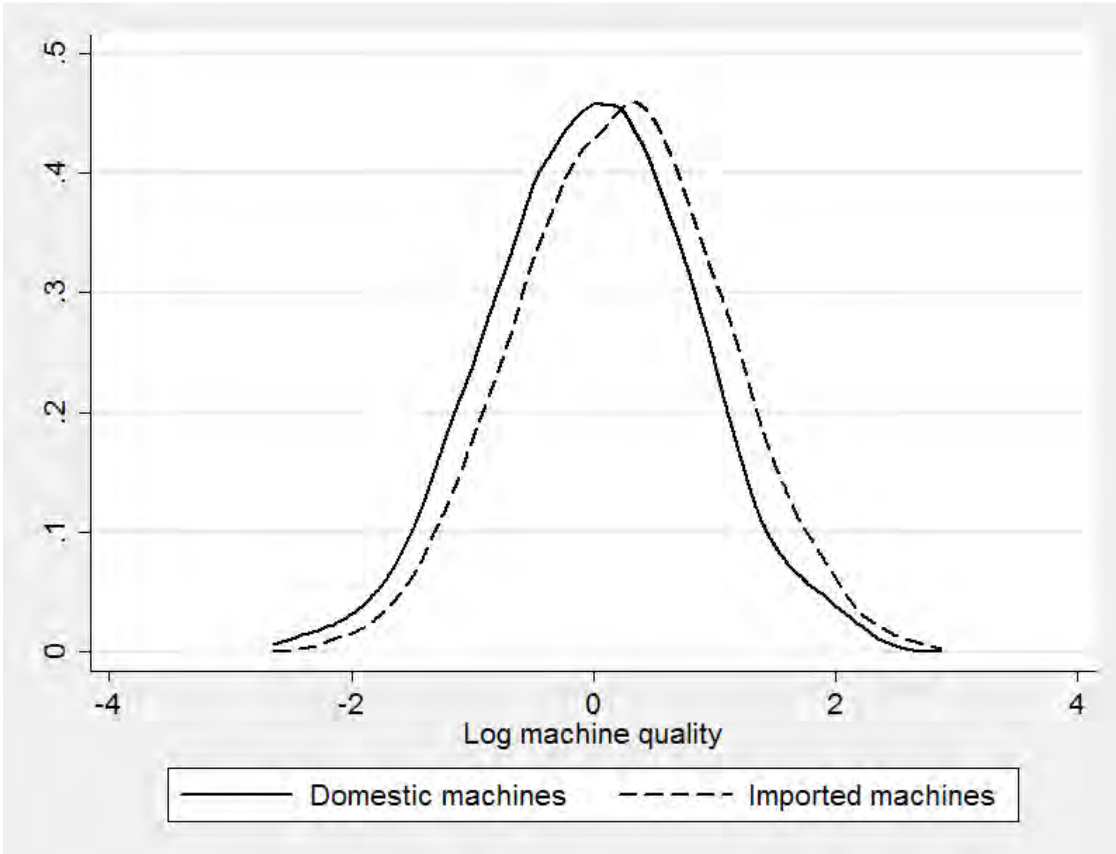

Figure 3: The estimated distribution of machine quality

are about 1, these can be understood as percentage differences.

Again, there is substantial variation of human capital within skill groups, and the interquartile ranges overlap significantly. The three schooling categories explain only 1.4 percent of the variation in human capital.

Figure 4 displays the estimated density functions for each skill group. Higher skill groups correspond to higher level of skill, but there is, again, considerable overlap across skill groups.

The wage of the lowest-skilled worker is estimated to be 0.53, about one half of the mean wage in the economy. That is, about one half of the wage of workers is pinned down by their outside option, and only the remaining one half is affected by their matching to machines.

The estimated wage gap between imported and domestic machines in this estimated economy is 12.4 percent. This is somewhat smaller than the actual wage gap of 16 percent. (The wage gap controlling for only occupation codes and hours of work is 0.162, when we get rid of those working at minimum wages, it is 0.161, the gap when using median regression is 0.165 and 0.163, respectively). Note that we have not targeted the wage gap directly.

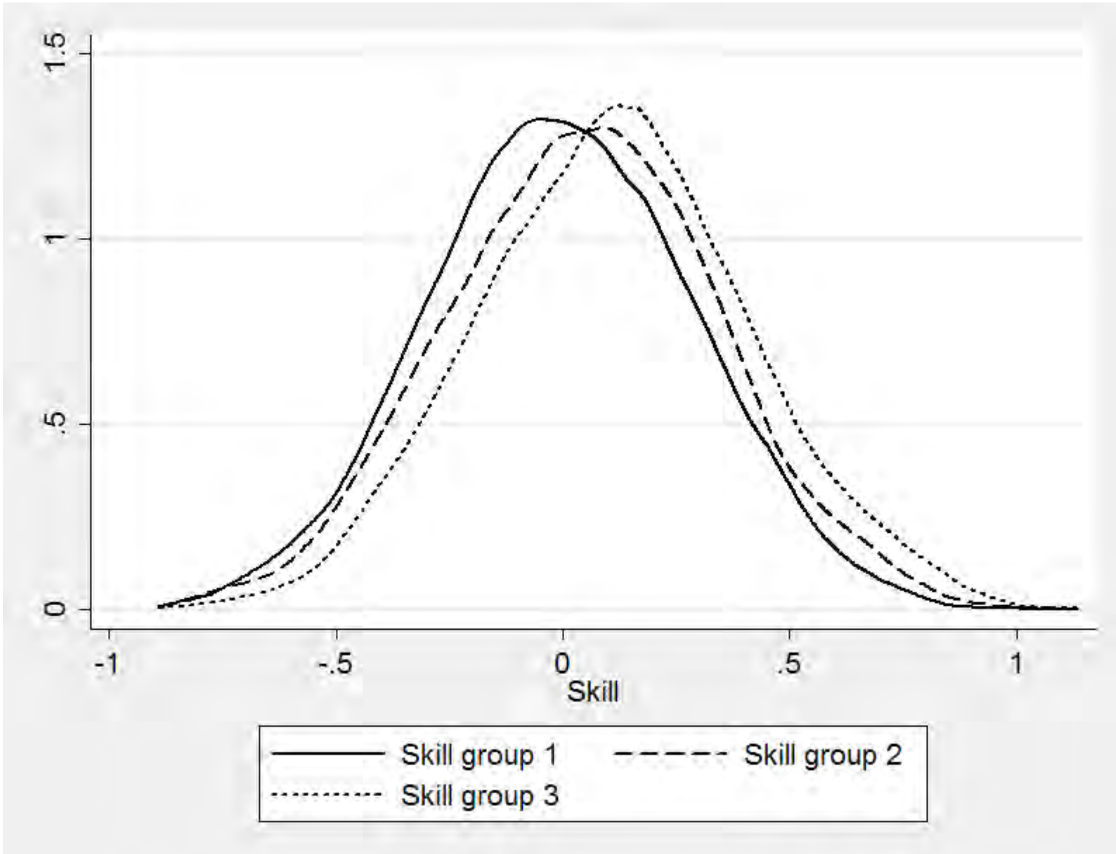

Figure 4: The estimated distribution of skill

Columns (2) through (4) report different specifications with only two groups of skills (no secondary, secondary schooling) and adding generic capital imports as a control. The estimated coefficients are similar, with the notable difference that the quality differential of imported machines drops by about one half if we control for capital imports. This can be interpreted two ways. One interpretation is that about one half of the effect estimated in columns (1) and (3) is due to firm selection. Another possibility is that because the import of generic capital and specific capital is highly correlated, once we control for generic capital imports, any measurement error in specific capital get magnified. Hence estimates in columns (3) and (4) may suffer from a stronger attenuation bias.

To evaluate the overall fit of the model, we plot actual wages against predicted wages in Figure 5. The dashed line is the 45-degree line, and the dots represent actual wage observations. Overall, the predicted and actual wages are highly correlated (the correlation coefficient is 0.99), with some notable deviations. First, the model misses the lower tail of the wage distribution entirely, where wages increase much more rapidly than predicted by the model. This may be the result of the minimum wage. Even though we have dropped all workers earning less than 60,000 forints, the minimum wage may affect the lower tail of the distribution to a greater extent by compressing wages there. Second, the model overpredicts the highest wages of machine operators. This may be the result of misspecified distributional assumptions.

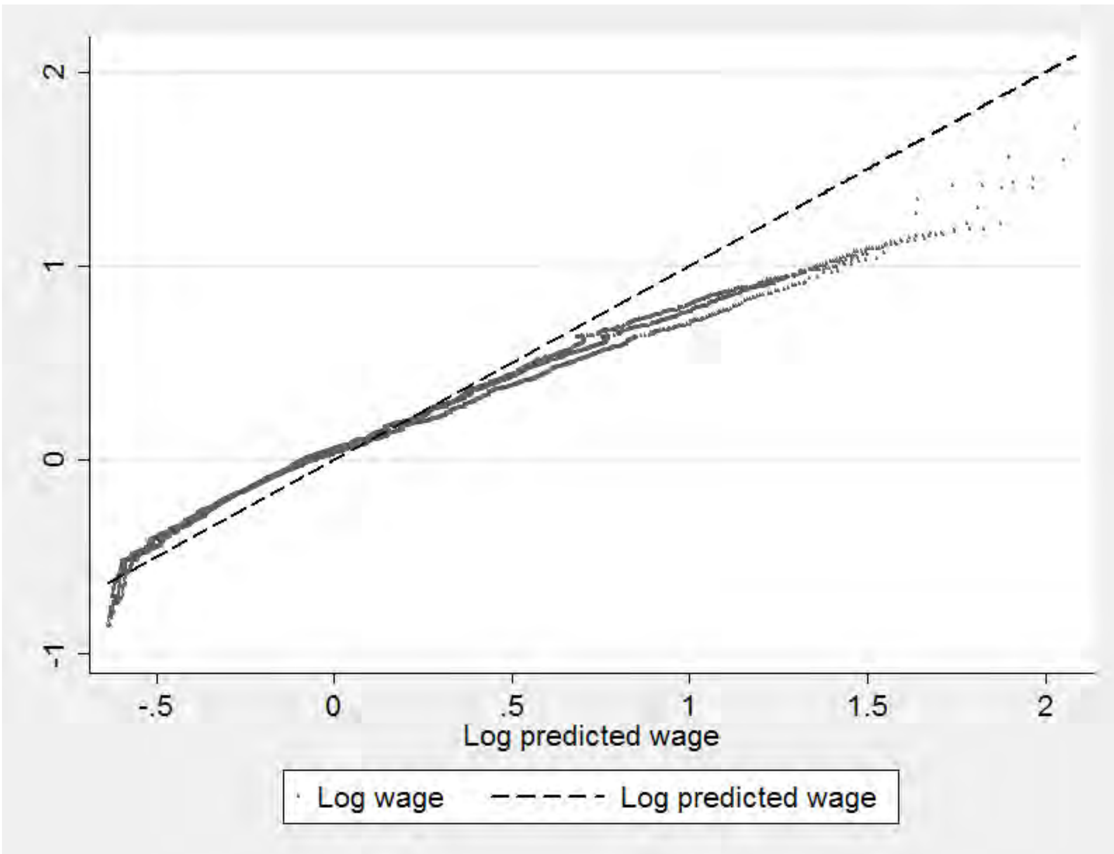

Figure 5: The fit of the structural model

6 The Effect of Imported Machines on Wages

To quantify the effect of imported machines on the level and inequality of wages, we conduct the following counterfactual experiment. We remove all imported machines from the economy by setting $\alpha_2 = 0$. We then replace them by domestic machines, the quality of each drawn from the same distribution as our estimates for G_1. This amounts to setting $\alpha_1 = 1$. All remaining parameter values are unchanged.

We ask a number of questions in this counterfactual economy. The results are summarized in Table 5. Column 1 reports several indicators for the estimated economy with $\alpha_2 = 0.34$. Column 2 reports these indicators for the counterfactual economy, with no imported machines. For completeness, we report the case of the counterfactual economy with only imported machines in column 3.

Table 5: Counterfactual experiments

Indicator	**Estimated economy**	**No imported machines**	**All imported machines**
Mean wage of importers	1.150	1.103	1.251
Mean skill of importers	0.093	0.093	0.093
Mean machine quality of importers	1.814	1.665	2.100
Mean wage	1.063	1.023	1.151
90/10 ratio of wages	2.587	2.479	2.792
75/25 ratio of wages	1.640	1.602	1.712

First, we measure the wages of those who work on imported machines *in the estimated economy*. All these workers will have to have switched to domestic machines in the economy with no imported machines. They earn 1.15 in the estimated economy, and about 4 percent less, 1.10, in the economy with no imported machines. This 4 percent is the direct effect of machine productivity. As the second row reports, the human capital of these workers are fixed. This implies that about one third of the 12 percent wage gap between imported and domestic machines is due to higher machine productivity, and the remaining two thirds are due to selection by (mostly unobserved) human capital.

How can the direct wage effect of imported machines be only 4 percent if they increase the returns to skill by 26 percent? First, recall that about one half of the wage is pinned down by outside options (w_0), and is not affected by machine productivity. Second, the 26 percent quality differential refers to the *average machine*. When we take away imported machines, their operators will be assigned to *better-than-average* domestic machines. This is because they have higher skills than average. Moving from the average imported machine to a better-than-average domestic machine entails a productivity loss that is less than 26 percent. In fact, we estimate it to be only 8 percent (row 3). The 8 percent productivity loss on one half of the wage rate yields the overall 4 percent wage drop when taking away imported machines.

Row 4 reports the mean wage in the economy, including workers on imported as well as domestic machines. It would be 3.8 percent lower without imported machines.

We then look at various measures of wage inequality. The 90/10 ratio is 2.59 in the estimated economy, and 2.48 in the simulated economy with no imported machines. The 75/25 ratio displays similar pattern, dropping from 1.64 in the estimated economy to 1.60 when we take away imported machines.

Column 3 of the table reports these indicators for the case when all machines are imported. Average machine quality and wages are higher, and there is more inequality in this case. The differences between all domestic (column 2) and all imported machines (column 3) are much larger than those between column 1 and 2. The intuition is that the selection mechanism that dampened the wage effects of imported machines are no longer present when the economy fully switches to imported machines. In this case, *everybody* will be working on a better-quality imported machine, not just the more skilled workers.

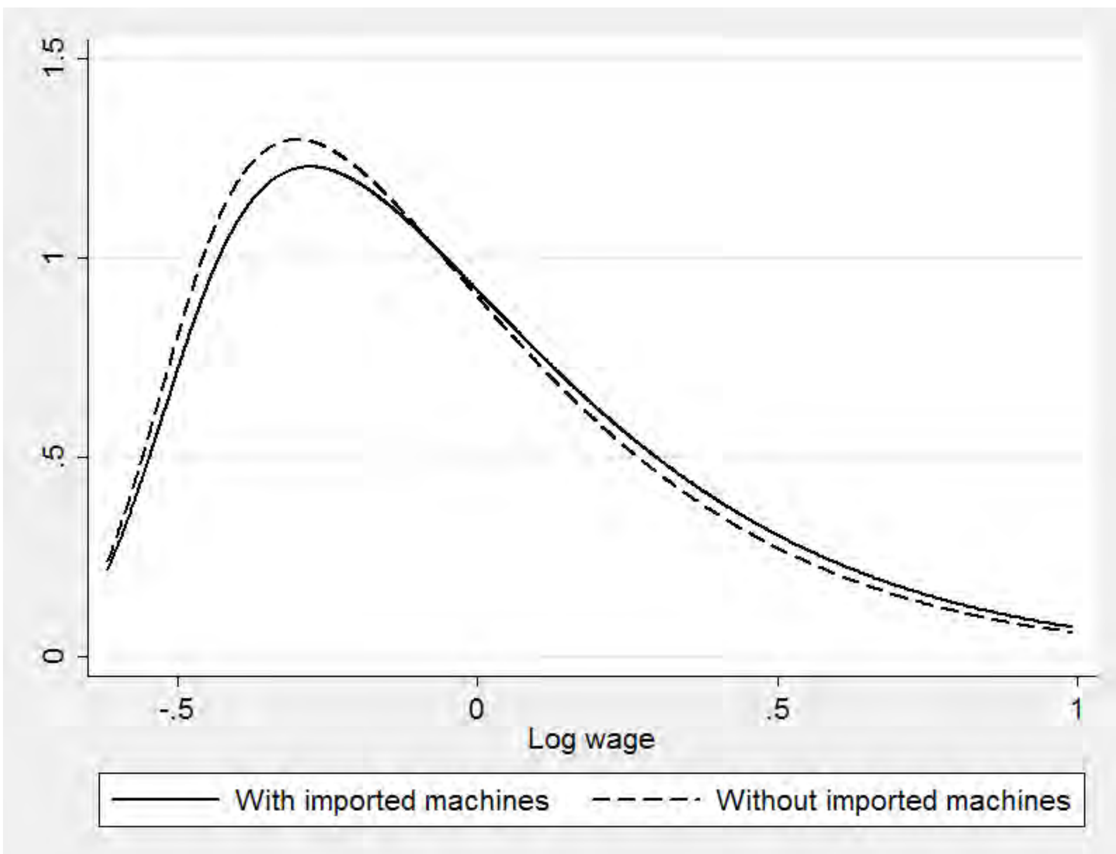

Figure 6: Wage density functions

Figure 6 displays the estimated density function for log wages in the two calibrated economies. The solid line corresponds to the benchmark economy, where 34 percent of machines are imported. The dashed line represents the counterfactual economy with no imported machines. Overall, we see that imported machines raise wages, in the sense that the density of wages is shifted to the right. Moreover, the wage density is more spread out with imported machines, reflecting increased inequality. However, most of the variation in wages comes from heterogeneity in human capital, so these changes are small relative to the overall wage distribution.

7 Conclusion

This paper estimated the effect of capital imports on the wages of a large, representative sample of Hungarian machine operators. Using linked employer-employee data and detailed product- and firm-level import data, we matched the precise occupation description of each worker to the equipment imported by their employers. We found that machine operators working on imported machines earn 16 percent more than those working on domestic machines. Estimating a structural assignment model of heterogenous workers and machines, we concluded that about one third of this wage gap is due to the higher returns to skill on imported machines, and two thirds are due to the higher skill of imported machine operators. Our structural estimates also suggest that imported machines contributed substantially to the increase in wage inequality in Hungary. Our results highlight a novel mechanism of skill-biased technical change.

References

Alfaro, Laura and Eliza Hammel, "Capital flows and capital goods," *Journal of International Economics*, May 2007, *72* (1), 128–150.

Altomonte, Carlo and Gábor Békés, "Trading Activities, Firms and Productivity," Working paper 2008.

Armenter, Roc and Miklos Koren, "Economies of Scale and the Size of Exporters," Working paper 2009.

Bernard, Andrew B, J Bradford Jensen, and Peter K Schott, "Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Good," in T. Dunne, J.B. Jensen, and M.J. Roberts, eds., *Producer Dynamics: New Evidence from Micro Data*, University of Chicago Press, 2009.

Bénabou, Roland, "Inequality and Growth," CEPR Discussion Papers 1450, C.E.P.R. Discussion Papers July 1996.

Bustos, Paula, "The Impact of Trade on Technology and Skill Upgrading: Evidence from Argentina," CREI working paper 2007.

Caselli, Francesco and Daniel J. Wilson, "Importing technology," *Journal of Monetary Economics*, January 2004, *51* (1), 1–32.

Costinot, Arnaud and Jonathan Vogel, "Matching and Inequality in the World Economy," NBER Working Papers 14672, National Bureau of Economic Research, Inc January 2009.

DiNardo, John E and Jorn-Steffen Pischke, "The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?," *The Quarterly Journal of Economics*, February 1997, *112* (1), 291–303.

Eaton, Jonathan and Samuel Kortum, "Trade in capital goods," *European Economic Review*, May 2001, *45*, 1195–1235.

Entorf, Horst, Michel Gollac, and Francis Kramarz, "New Technologies, Wages, and Worker Selection," *Journal of Labor Economics*, July 1999, *17* (3), 464–91.

Goldberg, Pinelopi Koujianou and Nina Pavcnik, "Distributional Effects of Globalization in Developing Countries," *Journal of Economic Literature*, March 2007, *45* (1), 39–82.

Halpern, László, Miklós Koren, and Ádám Szeidl, "Imported Inputs and Productivity," Working paper March 2009.

Heckman, James J and Bo E Honore, "The Empirical Content of the Roy Model," *Econometrica*, September 1990, *58* (5), 1121–49.

Jones, Charles I., "Intermediate Goods, Weak Links, and Superstars: A Theory of Economic Development," NBER Working Papers 13834, National Bureau of Economic Research, Inc March 2008.

Jovanovic, Boyan, "Vintage Capital and Inequality," *Review of Economic Dynamics*, April 1998, *1* (2), 497–530.

Kezdi, Gabor, "Two Phases of Labor Market Transition in Hungary: Inter-Sectoral Reallocation and Skill-Biased Technological Change," Budapest Working Papers on the Labour Market 0203, Institute of Economics, Hungarian Academy of Sciences July 2002.

Kremer, Michael, "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, August 1993, *108* (3), 551–75.

Krueger, Alan, "How Computers Have Changes the Wage Structure: Evidence From Microdata, 1984-89," *The Quarterly Journal of Economics*, 1993, *108* (1).

Roy, Andrew D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, June 1951, *3* (2), 135–146.

Verhoogen, Eric A., "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *The Quarterly Journal of Economics*, 05 2008, *123* (2), 489–530.

A Data Appendix

A.1 Matching machines to their operators

We match the 4-digit occupation code (FEOR) of machine operators to the 4-digit Harmonized System product code of capital goods. There are 74 FEOR codes involving the operation of a machine, including drivers of trucks and forklifts (we later wish to separate drivers). There are 105 HS codes describing machines or machine parts that we thought can be reasonably matched to workers. Some generic machines (e.g., computers) cannot be matched, so we dropped them from the sample. We then match each worker to at least one, potentially several machines that they can be working on. The average worker is matched with 4.5 machines, and the average machine is matched with 2.8 occupations.

Table 6 shows 20 randomly selected occupation/machine matches out of a total of 285. Some matches are more precise than others. For example, "Tobacco products machine operators" can be expected to work on "Machinery for preparing or making up tobacco." "Tugs and pusher craft" requires the work of "small boat steermen," but also of "Boatswains, chief engineers." "Metal working machine operators" are likely to work with "Files, pliers, metal-cutting shears," but probably also with other machines and tools.

Table 6: Some randomly selected matches of occupations and machines

pation ode	Description	Product code	Description
113	Tobacco products machine operators	8478	Machinery for preparing or making up tobacco, not elsewhere specified in this chapter
353	Metro drivers	8603	Self-ppld railway/tramway coache, vans & truck o/t those of heading no8
313	Plant protection machine operators	8436	Agricultural, hortic, forest, bee keeping mchy; poultry incubator etc
364	Small boat steermen	8904	Tugs and pusher craft
362	Boatswains, chief engineers	8904	Tugs and pusher craft
351	Locomotive engine and train drivers	8602	Rail locomotives, not elsewhere specified; locomotive tenders
319	Agricultural and forestry mobile-plant drivers, operators not elsewh	8432	Agricultural, hortic/forest machinery for soil prep/cultivation; lawn-rollers
240	Packaging machine operators	8422	Dish washg mach; machinery for clean/drying/ fil/clo/etc; machinery for aerating
331	Scavengery machine operators and drivers	9104	Instrument panel clock & clock of similar type for veh, aircraft, vessel,
291	Boiler operators (licensed boilermen)	7322	Radiators & parts thereof, i/s; air heaters & hot air distributors
293	Agricultural machine operators, mechanics	8201	Hand tools of a kind used in agriculture horticulture or forestry
322	Groundwork machine operators	8429	Self-propelld bulldozer/angledozer/ grader/excavator/shovel loader,
361	Bargemen	9104	Instrument panel clock & clock of similar type for veh, aircraft, vessel,
351	Locomotive engine and train drivers	8408	Compression-ignition int combu piston engine (diesel or semi-diesel eng)
192	Metal working machine operators	8203	Files, pliers(incl cut pliers), pincers, met cut shears, etc & similar ha
364	Small boat steermen	8905	Light vessel, dredger; floating dock; floating/submersible drill platforms
313	Plant protection machine operators	8432	Agricultural, hortic/forest machinery for soil prep/cultivation; lawn-rollers
311	Agricultural engine drivers and operators	8434	Milking machines and dairy machinery
329	Construction machine operators not elsewhere classified	8429	Self-propelld bulldozer/angledozer/ grader/excavator/shovel loader,
144	Concrete building block machine operators	8424	Mechanical appliance for proj/dispersing/spray liq/pdr; sand blasting machinery,