Replication : The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation (2022)

Mathias Dachert

November 6, 2022

1 Introduction

Where are the differencies among nations coming from? This question is of major importance in economics and for developpement studies. The theories of growth brought explanations about what could make a country richer than an other one as for Lucas and Romer (Romer (1990), Lucas (1988)) who introduced the Endogenous Groth model and the effect of accumulation of human capital on economic developpment. But later in the litterature Caselli and Ii (2006) added a very intuitive feature to the model: imperfect substituability between skilled and unskilled labor. With this simple specification and reameasuring for skill bias accross nations they found that technology bias plays a more improntant role than we thought. Following this reasoning the theory is that rich countries use their skills better because they have an industrial structures and technologies suited for skilled workers. However the idea that human capital differences are more striking come back with Jones (2014) and the new concept of Human Capital Aggregator. This aggregator function is defining the relative efficiency of the skilled workers. In these models human capital plays a more important role. The explanation given is that the quality of education in rich countries widens the gap in effeciency between low and high skilled workers. These two adverse ideas in developpement accounting are called the "relative technology" interpretation and the "relative human capital" interpretation.
There comes the paper of Frederico Rossi (Rossi (2022)): How can we find a way to evaluate both theories? Some answers were already given by Okoye (2016), but what Frederico Rossi brings to the litterature are: The usage of micro-data to better measure the relative skill efficiency and an interesting identification strategy using migration as a counterfactual.

2 Summary

The first stone of Rossi (2022) is a measurement work using micro Data. This paper uses a standard model and brings the following euqation to the data.

$$\frac{w_{H,c}}{w_{L,c}} = \frac{A_{H,c}Q_{H,c}}{A_{L,c}Q_{L,c}}\frac{F_H(.)}{F_L(.)} \tag{1}$$

Where the first fraction is the wage ratio between high and low skilled workers. We do observe it in the data but the main interest is the first fraction on the right hand side. It's the relative efficiency between high and low skilled workers. The last term is the relative efficiency of all the other factor of efficiency like technology or environement. What we want to know is how much of the wage ratio is actually explained by the relative efficiency of human capital ($\frac{A_{H,c}Q_{H,c}}{A_{L,c}Q_{L,c}}$). This is where the micro-data used by this paper is powerful, the relative efficiency of Labor units is directly observed, having all the information on the sector and wages of workers. Then we can compare with the wage ratio and infer the weight of Human Capital and the environement.

The second strategy is to use immigrants workers as counterfactual. In which sense immigrants represent a counterfactual? This paper uses the fact that immigrants were educated in another country, possibility a lower income country, but are working in the same country as the natives who received a better education. The identification strategy here is to sort by country of origin and see if there is a relation between being from a low income country and having a low labor efficiency.

3 Replication of the results

The Data

The Data wasn't fully provided by the author in his replication package. The IPUMS international noa (2022)micro-data were missing. In order to replicate I had to import them and given the capacity of my computer I could not use the Brazil Dataset. Additionally, the Dataset on India wasn't available in the IPUMS website. However all the other Dataset (and of course Ruggles et al. (2022)) were easy of access or provided by the author. I nonetheless have to note that the Data for the US was too eavy to clean myself so I had to use the final Dataset provided by the Author.

The code

The code was working correctly but running all the do files takes a long time due to the dimension of the Datasets. The author mentionned that with a 16GB RAM computer he was able to compute all the code in 5 hours. With my 8GB RAM and consider all the adjustement for it not to shut down it was some days. Despite the technical aspect of it I was able to compute the main result that I will present.

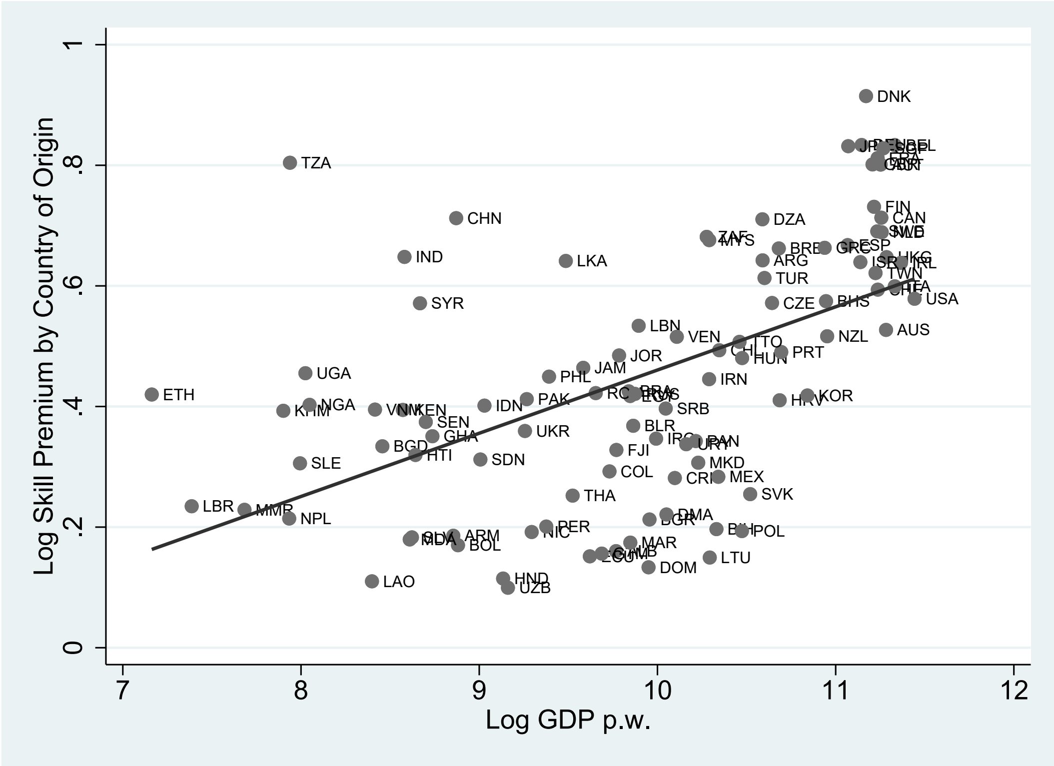
The results

Table 1: Skill Premium, Supply, and Efficiency across Countries

	Baseline			No hours	All working age	Traditional
country	w_H/w_L	$\tilde{H}/\tilde{L}$	$(A_HQ_H)/(A_LQ_L)$	$(A_HQ_H)/(A_LQ_L)$	$(A_HQ_H)/(A_LQ_L)$	$(A_HQ_H)/(A_LQ_L)$
Indonesia	1.957	0.070	0.003	0.004	0.009	0.006
Jamaica	2.969	0.067	0.010	0.011	0.011	0.003
Venezuela	2.490	0.257	0.089	0.132	0.152	0.055
Uruguay	2.218	0.363	0.126	0.189	0.225	0.260
Panama	2.262	0.313	0.099	0.123	0.119	0.077
Mexico	2.205	0.227	0.049	0.069	0.070	0.040
Trinidad and Tobago	2.746	0.100	0.018	0.022	0.024	0.009
Israel	1.606	0.596	0.129	0.155	0.109	0.156
Canada	1.508	1.539	0.711	0.825	0.928	1.628
United States	1.802	1.397	1.000	1.000	1.000	1.000
Elasticity wrt GDP p.w.	-0.174	1.308	2.093	2.034	1.785	2.281
	[0.084]	[0.247]	[0.387]	[0.411]	[0.439]	[0.576]

The main table is about the measurement contribution of the author. As we mentionned the paper using micro Data is able to compute the statistics of interest, wage ratio and relative efficiency in a more consistent way than previous results. I find very similar (identical most of time) results even if I couldn't include India and Brazil for technical reason as said before. The only difference with the results of the paper are the elasticities which are much higher in my case (2.093 against 1.408 for the relative efficiency of skilled labor) We can explain this difference by the absence of Brazil and India creating more extreme disparities among nations and also much more observations relatively in the high income countries. The interpretation of these findings are still the same the relative skill efficiency doesn't explain the major part of the skill premium. And the interpration of the elasticities with respect to GDP per capita are even more strong. This relative efficiency is higher when GDP increases.

This is where the second part of the paper comes in. The first result is a graphic which might be counterintuitive in the interpretation. Even if graphic 1 shows a positive relation betwen GDP and skill premium by country of origine it is actually very low compared to the explanation power we usually give to human capital differences.



This graph is very similar to one found in the paper. To understand better this question of the immigrants as counterfactual we have to look at the numerical results.

Table 2: Relative Human Capital across Countries

	Broad sample				Microdata sample			
		θ_Q/θ_{AQ}				θ_Q/θ_{AQ}		
	θ_Q	$\sigma = 1.5$	$\sigma = 1.3$	$\sigma = 2$	θ_Q	$\sigma = 1.5$	$\sigma = 1.3$	$\sigma = 2$
US Immigrants	0.105	.095	.057	.189	0.043		.018	.011
	(0.016)				(0.048)			
Baseline (Pooled)	0.106	.096	.058	.192	0.075		.031	.019
	(0.016)				(0.041)			
Bilateral Controls	0.448	.404	.243	.809	1.134		.474	.281
	(0.101)				(0.473)			

The entries are a ratio betwen two elasticities, the elasticity with respect to human capital and the total relative efficiency elasticity. We don't have the same results for probabily two reasons, we excluded India and Brazil and due to technical issues I was not able to run the program many times and so I couldn't double check my results. In anyway the results are similar, and what is important here is that among immigrants human capital is approximatically 10 percent of relative efficiency. Which is the conclusion of Frederico Rossi, Human capital seem to not play the most important role in relative skill efficiency. But we could question the way Frederico Rossi uses the natural experiment of immigration to construct a ceteris paribus reasoning.

4 Extension

The question we ask in this extension is wether immigrants represent a good counterfactual. The reasoning here is the same as the balance check in RCT. Of course to compute his results Frederico Rossi didn't have to assume any randomness but he stills augue that gender for example isn't an issue to his work. "Moreover, I leverage the individual-level information available in my dataset to show that, to a large extent,cross-country gaps in relative skill efficiency are not driven by differences in sectoral composition, in the incidence of self-employment or in the returns to other observable characteristics, such as gender and experience." Rossi (2022). Even he controlled all regressions with the observable characterestics nothing stop us from looking a bit into it. We start by checking a simple regression on US immigrants.

Table 3: Balance check immigrants

	(1) immigrant
sex	-0.00162***
	(-12.88)
educ	-0.0151***
	(-568.09)
_cons	0.239***
	(892.12)
N	28830806

t statistics in parentheses

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

To understand the coefficient on gender it's important to say that according to IPUMS coding Male is 1 and Female is 2. The conclusion here if it is not large there is indeed a significant relation between being an immigrant and beeing a male. The regression on education was to make sure immigrants have indeed less educational attainment as argued by the author. The following replication exercise I will do recompute the results without the female population in the samples to see if we find the same trend. Of course in this exercise again the size of the Data was an issue for me and my computer. So I was not able to compute the results for US immigrants but I could do it for all other countries' immigrants.

Table 4: Skill Premium, Supply, and Efficiency across Countries among Male

	Baseline			No hours	All working age	Traditional
country	w_H/w_L	$\tilde{H}/\tilde{L}$	$(A_HQ_H)/(A_LQ_L)$	$(A_HQ_H)/(A_LQ_L)$	$(A_HQ_H)/(A_LQ_L)$	$(A_HQ_H)/(A_LQ_L)$
Indonesia	1.957	0.070	0.003	0.004	0.009	0.006
Jamaica	2.969	0.067	0.010	0.011	0.011	0.003
Venezuela	2.490	0.257	0.089	0.132	0.152	0.055
Uruguay	2.218	0.363	0.126	0.189	0.225	0.260
Panama	2.262	0.313	0.099	0.123	0.119	0.077
Mexico	2.205	0.227	0.049	0.069	0.070	0.040
Trinidad and Tobago	2.746	0.100	0.018	0.022	0.024	0.009
Israel	1.606	0.596	0.129	0.155	0.109	0.156
Canada	1.508	1.539	0.711	0.825	0.928	1.628
United States	1.802	1.397	1.000	1.000	1.000	1.000
Elasticity wrt GDP p.w.	-0.174	1.308	2.093	2.034	1.785	2.281
	[0.084]	[0.247]	[0.387]	[0.411]	[0.439]	[0.576]

We find very similar, almost identical, results on the relative efficiency and the wage ratio which shows that the statistics are not influenced by gender in any ways.

Table 5: Relative Human Capital across Countries among Male

	Broad sample				Microdata sample			
		θ_Q/θ_{AQ}				θ_Q/θ_{AQ}		
	θ_Q	$\sigma = 1.5$	$\sigma = 1.3$	$\sigma = 2$	θ_Q	$\sigma = 1.5$	$\sigma = 1.3$	$\sigma = 2$
US Immigrants	0.105	.095	.057	.189	0.043		.021	.012
	(0.016)				(0.048)			
Baseline (Pooled)	0.105	.095	.057	.19	0.063		.03	.018

In the conterfactual exercise we find again very similar results which conclude on this extension. Indeed, gender doesn't seem to play a big role in explaining relative skill efficiency among immigrants.

5 Conclusion

Through this replication I have found that the conclusions of the paper Rossi (2022) are reproducible, to a certain extend, and robust to simple balance check extension. Of course this replication could go much further in all its dimensions. The technical constraint joined with the time contraint led this work to be less deep than planned. In this conclusion, I would like for later continuation give some ideas for a follow up. I would be interesting to get deeper in the model used to go the Data. Even if this part of the paper is weel documented. Maybe we could introduce an international dimension to the model in example in an population growth. Finally a simple extension one could add is to add more countries to the analysis or add more variables as controls. However I didn't do it for the same technical contraints mentionned before.

References

IPUMS International: Version 7.3, 2022. URL `https://international.ipums.org/international/index.shtml`.

Francesco Caselli and Wilbur John Coleman Ii. The World Technology Frontier. *THE AMERICAN ECONOMIC REVIEW*, 96(3):24, 2006.

Benjamin F. Jones. The Human Capital Stock: A Generalized Approach. *American Economic Review*, 104(11):3752–3777, November 2014. ISSN 0002-8282. doi: 10.1257/aer.104.11.3752. URL `https://www.aeaweb.org/articles?id=10.1257/aer.104.11.3752`.

Robert E. Lucas. On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42, July 1988. ISSN 0304-3932. doi: 10.1016/0304-3932(88)90168-7. URL `https://www.sciencedirect.com/science/article/pii/0304393288901687`.

Dozie Okoye. Appropriate Technology and Income Differences. *International Economic Review*, 57(3):955–996, 2016. ISSN 0020-6598. URL `http://www.jstor.org/stable/44075375`. Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University].

Paul M. Romer. Human capital and growth: Theory and evidence. *Carnegie-Rochester Conference Series on Public Policy*, 32:251–286, March 1990. ISSN 0167-2231. doi: 10.1016/0167-2231(90)90028-J. URL `https://www.sciencedirect.com/science/article/pii/016722319090028J`.

Federico Rossi. The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation. *American Economic Review*, 112(1):235–266, January 2022. ISSN 0002-8282. doi: 10.1257/aer.20191852. URL `https://pubs.aeaweb.org/doi/10.1257/aer.20191852`.

Steven Ruggles, Sarah Flood, Ronald Goeken, Megan Schouweiler, and Matthew Sobek. IPUMS USA: Version 12.0, 2022. URL https://usa.ipums.org. Version Number: 12.0 Type: dataset.