Jacob Watt-Morse
ECON 459: Economics of Institutions
Professor Ashraf
5/23/2016

Geographic Oddities: Untangling Geographic Controls from On the Origin of Gender Roles

Introduction

Alberto Alesina, Paola Giuliano and Nathan Nunn set out to determine the cause of the considerable heterogeneity in gender roles around the world today. They work through Ester Boserup's hypothesis that gender norms established by the adoption of plough agriculture persist today. They access this hypothesis in several different ways. They test if plough adoption affected preindustrial agricultural labor distribution. They regress ethnic ancestral plough use against country level outcomes. They test their findings against several alternate theories, run an instrumental variable regression to eliminate reverse causality and use data on immigrants to determine if these outcomes are the results of cultural norms or institutions. While many of results support this conjecture, a flaw in the regression using World Value Survey (WVS) data and the reduced form result of their instrumental variable regressions indicate there is a potential omitted variable. I then reevaluate their cross country regressions with two new variables, present day tropical climate and Malaria prevalence rate. These results show this omitted variable is highly correlated with present day tropical climate but unrelated to tropical diseases.

Paper Summary

The hypothesis these authors are testing was originally advanced by Ester Boserup in her 1970 paper *Women in Economic Development.* She noted men had a considerable comparative advantage in plough intensive agriculture because the plough requires a lot of brute upper body strength. Responding to this advantage, men came to dominate pre-industrial agricultural tasks. She argues these pre-industrial workforce division led to gender roles that persist to present day; effecting the distribution of labor and roles women play in other economic and political positions.[1]

The authors use data from the *Ethnographic Atlas* to measure preindustrial plough use. This data set, compiled by George Peter Murdock, has information on 1,265 ethnic groups. These observations are the earliest reliable studies of these groups. In other words the data set is a compilation of the first satisfactory examinations of each individual ethnic group. Included in this data set is a variable that indicates if the plough was absent, existed but was not aboriginal or existed aboriginally. From this measurement the authors create an indicator variable for plough use which takes on the value 1 if the plough was present in pre-industrial society.[2]

The first part of Boserup's Hypothesis they examine is the effect of plough adoption on female participation in preindustrial agriculture. To do so they regress the plough indicator

[1] Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." Quarterly Journal of Economics, 128(2): 470-71
[2] Ibid. 478

against a scaled variable from the *Ethnographic Atlas* that measures the gender composition of the preindustrial agricultural sector. They include a vector of controls for ethnic economic development, the presence of domesticated animals, as well as tropical climate and agricultural suitability measures from the *Global Agro-Ecological Zones*. They find that presence of the plough had a large and significantly negative effect on female participation in agriculture. They supplement these findings with data from the *Standard Cross-Cultural Sample* (SCCS) which includes observations of specific tasks. While the sample is smaller the authors claim it is representative of every ethnicities. They find similar results from these regressions. Plough use is associated with a significant decrease in female participation in all agricultural related tasks. The authors argue that since the plough is only used in soil preparation, detecting an effect in every tasks indicates gender norms about work were engendered by the presence of the plough. Each of these results indicate the first step in Boserup's causal chain is highly probable.[3]

Next the authors run a present day cross country regression to examine if plough use is correlated with female labor force participation. Instead of using a plough indicator variable, they create a proportion of ethnic ancestry that used the plough for a given country. They map present day populations with ancestral ethnic groups by using data from the *Ethnologies: Languages of the World* data set. They link current day languages to their ethnic ancestors and aggregate these totals over a whole country to get a proportion of the population with ancestors that used the plough. Three different indicators of gender roles, female labor force participation in 2000 (FLFP), share of firms with female owners and the share of political positions held by women, are then regressed against that measure. The authors include the same ethnic control vector from before and a control for income and income squared. Further half the regressions have continent fixed effects. The authors find a significant negative relationship between plough ancestry and their indicator variables. A one standard deviation increase in ancestral plough use is associated with a nearly 6 percent decrease in FLFP. These results show that Boserup's hypothesis captured a strong correlation if not causation.[4]

The authors then turn to sub-national data in hopes of getting more precise estimates. First they analyze data from the World Value Survey (WVS) to measure if ancestral plough use effects both the labor market and underlying gender norms. The survey does not always ask individual respondents for their ethnicity, so the authors use the proportion of plough ancestry within a district as a measure of individual plough use ancestry. They regress FLFP, and two questions from the survey: "When jobs are scarce, men should have more right to a job then women" and "On the whole men make better political leaders then women do," against this district measure. They include the same controls as the last regression except half the regressions contain continent fixed effects while the other half contain country fixed effects. The results for the continent fixed effects support their results. This is not the case for the country fixed effects regressions. The point estimate on ancestral plough use (the authors call it traditional plough use) in the FLFP regression is almost zero and statistically insignificant while the coefficient in the job scarcity regression is only significant at the 10% level. The authors claim these weaker findings are a result of attenuation bias which arises from the inability to observe individual respondents ancestry. I will return to these concerns later in the critique.[5]

---

[3] Ibid. 480-83
[4] Ibid. 488-496
[5] Ibid. 501-505

To address this attenuations bias the authors regress individual outcomes against plough ancestry using data from the *IPUMS-International Census.* Respondents report their ethnicity so plough ancestry is directly measured rather than estimated. These regressions include the same controls as before. They find a significant and negative relationship between ethnic ancestral plough use and FLFP. The data set only includes information on 8 countries, none of which were in North American or Europe, and thus may not be representative of the whole sample.[6]

The authors next try to prove causality using three strategies. First, they repeat their cross country analysis but include a variety of controls that capture alternative theories that could be correlated with ancestral plough use. Frist, is pre-industrial intensive agricultural practices to check if plough use is simply picking up the effects of intensive agriculture. Second, is a variety of familial structural measures because their plough variable could simply capture difference in communal hierarchies. Third, is the year of observation of different ethnicities to capture any possible measurement error from the *Ethnological Atlas*. Fourth is warfare and political indicators like the prevalence of communism. Fifth is the size of manufacturing, oil and service sectors as each sector has different gender composition. Sixth is a control for the prevalence of certain religions, as religion was highly correlated with plough use. Lastly they include all these controls in a single regression. The coefficient on their traditional plough use variable remains highly significant and consistent through all of these controls. These results show ancestral plough use is not capturing these alternative theories. However, these regression do not eliminate the possibility an omitted variable could work against their findings.[7]

The second strategy is a 2SLS regression which examines Fredrick Pryor's hypothesis that plough adoption was a process endogenously determined by geographic suitability. By determining the origins of plough use, they demonstrate that plough adoption was not caused by gender norms. Their results support Pryor's hypothesis that plough climatic suitability determined plough use as instrumented plough use does has significant effect on current day gender roles. One problem with their result is negative plough suitability is positively correlated with FLFP in reduced form, even though there is no correlation between negative plough suitability and ancestral plough use. This indicates there is some exclusion violation and this violation is likely some geographic conditions.

Lastly the authors seek to exhibit that ancestral plough use effects these outcomes through the transmission of cultural norms and not the construction of sexist institutions. They do so by using the "natural experiment" of immigration. Observing immigrants allows them to examine variation in culture while keeping institutions constant. First they observe if the proportion of ancestral plough use in the emigrated country had an effect on FLFP. They conclude that it does and that the "transmission of internal norms explains about 35-50% of the total effect measured earlier."[8] Next they regress the responses of European immigrants to the job scarcity question asked in the WVS against ancestral plough in the emigrated country. They again find that about 35-50% of the total effect measure earlier comes from the transmission of norms. These results should be taken with a grain of a salt as they are subject to the same

[6] Ibid. 505-506
[7] Ibid. 508-514
[8] Ibid. 521-524

attenuation bias observed in the WVS regressions. Further, that bias is compounded by the fact that immigration is not random and these immigrants may not represented the ethnic majorities in their country.[9]

In summary, Alesina et al. test Boserup's hypothesis in a variety of ways with a varying degree of success and rigor. Their results are sometimes surprisingly consistent and economically significant. However, at other times their results fall flat and their some of their empirical strategies have several mechanical deficiencies.

Analyzing the Failings

I now return to the findings that do not support Boserup's hypothesis. Primarily, the regressions using WVS data that include country fixed effects. The traditional plough use coefficient in the regressions using FLFP as the dependent variable is insignificant and has negligible magnitude (-.002). This indicates that their plough variable may be capturing a cause of gender role heterogeneity that is highly correlated with aggregate country traditional plough use. This omitted cause cannot be one of the many controls they use in the latter section of the paper, as that control would have had a noticeable effect on the plough use coefficient. The authors overlook some condition that is highly correlated with overall country level plough use that may explain the result they find.

Another indicator that the authors omitted some causal variable is apparent in their instrumental variable regressions. The negative geographic plough suitability variable is positively correlated with present day FLPF, even though it is not statistically significantly correlated with ancestral plough use. This shows there is some geographic cause of gender roles that doesn't work through plough adoption. Taken together these two regression indicate there is some geographic condition that is highly correlated with country level ancestral plough use which has an effect on modern day gender roles.

This deep determinant is likely something that is considerable different between the west and the developing world. Figure II from their paper (print below) exhibits how there is not much within continent variation in ancestral plough use but strong differences between continents. This is best exhibited by the difference between sub-Saharan African and Europe. The authors notice this concern in their introduction:

> There is little variation within Europe and within sub-Saharan Africa. Therefore, there is the concern that our estimates are simply driven by broad differences between regions like Europe and Africa.[10]

While they do their best to address this by using individual level observations, they note they cannot completely eliminate this possible flaw in their results. When combined with the regression faults noted above this mechanical shortcoming indicates there is geographic condition that is highly correlated with country level ancestral plough use and differs considerable between developing and developed countries that effects modern day gender roles.

---

[9] Ibid. 524-526.
[10] Ibid. 486

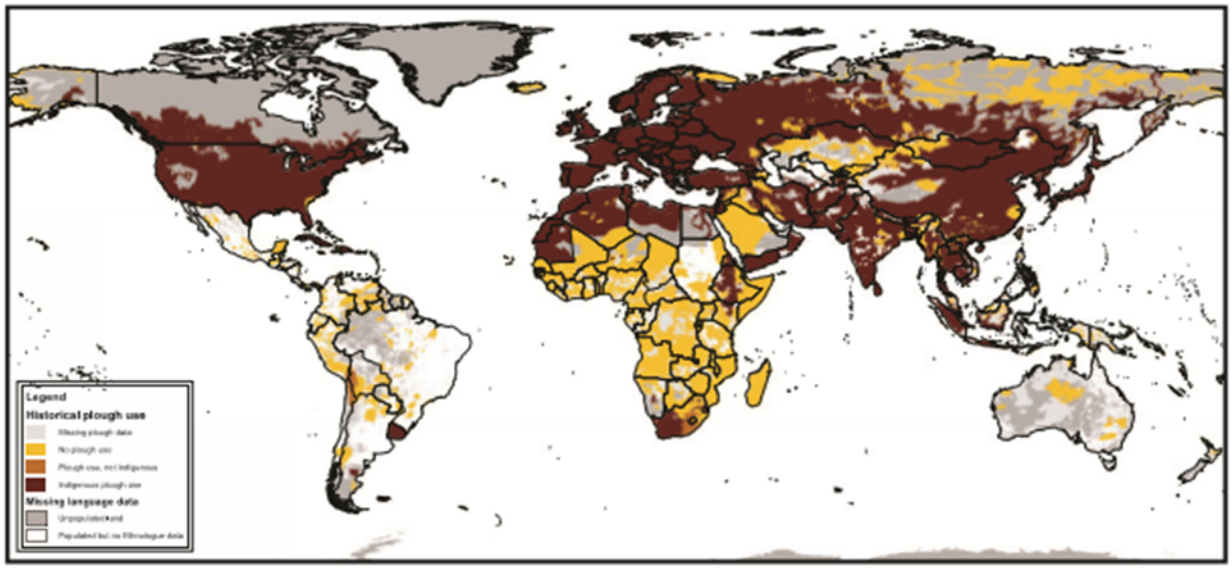

FIGURE II
Traditional Plough Use among the Ethnic/Language Groups Globally

Empirical Strategy: Testing Climate

I test if their results hold with the inclusion of two known geographic differences between the developed and developing world: present day population weighted tropical climate and malaria prevalence. Both of these measures were omitted form their length list of controls. I supplement the cross country regression data from Alesina et al. with cross country data from Ashraf et al. [2013] which includes both of these measures. Each variable is highly correlated with negative plough suitability (topical climate .602, malaria .577). Hence, these could also cause the reduced form result from the IV regressions negative plough suitability.

While Alesina et al. control for tropical climate, the control they use accounts for ethnic group's residency in tropical climates. Therefore, any effect of present day tropical climate on gender roles is not necessarily capture by this variable. The correlation between ethnic tropical climate and present day tropical climate (pictured below) demonstrates that Alesina et al. tropical variable is not representative of modern day climate. Further their tropical control may not even capture the full effect of tropical climate on ethnic groups. They control for fraction of ethnic land in tropical zones. This measure does not capture how many people were living in tropical conditions and therefore may skew their results.

To address these concerns I simply repeat their initial cross country regressions with the inclusion of these two variables. The independent variable of interest is the proportion of the country with Plough ancestry. The dependent variable of interest is female labor force participation in 2000. The same controls for ethnic conditions and present day income that Alesina et al. use are included as well.

Results

My findings are presented in table 1 printed below. Column 1 replicates their findings from table 4, column 1. Column 2 replaces ethnic tropical climate with present day tropical climate. The statistical significance on traditional plough use drops from the 1% level to the 10% level, a noticeable drop in explanatory power. This indicates that traditional plough use was capturing some variation in present day tropical climate. Column 3 includes the malaria control, which is insignificant. The coefficient on traditional plough use is significant and similar to the point estimate in column 1. This shows that the malaria prevalence is not correlated with FLPF.

Column 4 includes both present day tropical climate and malaria prevalence rate. Including both completely eliminates the significance of traditional plough use while present day tropical climate remains highly significant. This again exhibits traditional plough use was actually capturing the effect of present day tropical climate. The malaria point estimate moves closer to zero, further demonstrating that Malaria has no strong correlation with FLFP.

Column 5 includes both ethnic and present day geographic controls as well as a control for malaria prevalence. While the traditional plough use variable is significant it has both a lower t-statistic and smaller point estimate then the original regression. This regression, while to a lesser extent, demonstrates that traditional plough use captured much of present day tropical climates effect on female gender roles.

Table 1
Dependent Variable: Female Labor Force Participation in 2000

	(1)	(2)	(3)	(4)	(5)
Traditional Plough Use	-12.40***	-7.091*	-11.06***	-6.671	-7.794**
	(2.964)	(4.085)	(3.491)	(4.357)	(3.935)
Historical Controls					
Agricultural suitability	6.073	13.14***	6.061	13.26***	8.805**
	(3.696)	(3.982)	(3.847)	(4.005)	(3.948)
Ethnic Tropical Climate	-9.718***		-10.84***		-13.60***
	(2.487)		(2.662)		(2.980)
Large animals	-2.015	0.689	-5.312	0.0689	-2.633
	(5.372)	(5.469)	(5.804)	(5.968)	(6.417)
Political hierarchies	0.779	1.807	1.467	1.865	1.728
	(1.515)	(1.492)	(1.526)	(1.506)	(1.411)
Economic complexity	1.157	-0.293	0.822	-0.338	-0.572
	(0.793)	(0.999)	(0.745)	(1.001)	(0.871)
Contemporary Controls					
ln income in 2000	-34.61***	-35.06***	-29.41***	-33.96***	-28.99***
	(6.528)	(6.478)	(7.428)	(7.562)	(7.317)
ln income squared in 2000	2.038***	2.160***	1.751***	2.100***	1.755***
	(0.406)	(0.410)	(0.454)	(0.470)	(0.449)
Added Controls					
Present Day Tropical Clim.		9.506**		9.318**	12.24***
		(4.249)		(4.316)	(4.019)
Malaria Prevalence Rate			6.684	1.467	5.135
			(4.635)	(4.500)	(4.335)
Constant	193.8***	175.7***	172.6***	171.0***	169.0***
	(26.42)	(25.72)	(30.32)	(30.09)	(29.87)
Observations	165	150	152	150	150
R-squared	0.399	0.409	0.422	0.410	0.480

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Discussion: What's going on with these climate variables?

Table 1 includes a couple of interesting findings that may help explain what these authors overlooked. First off, these regressions demonstrate that (much to Jeffrey Sachs' chagrin) malaria appears to have no effect on FLFP. Malaria acts as an imperfect proxy for other tropical diseases that affect both humans and livestock. For instance, the Tsetse fly is prevalent in areas with high rates of malaria. While not entirely conclusive, these results cast considerable doubt on the theory that tropical disease somehow effect gender roles.

On the other hand these results raise a lot of questions about how Alesina et al. handle tropical climate. Not only does including for present day tropical climate reduce the significance of the regression results it also behaves unexpectedly with other variables. The two geographic controls have different signs, with present day geography being positively correlated and ethnic geography being negatively correlated with current day gender roles. Furthermore when both are included, the magnitudes on both coefficients increases dramatically.

This finding could results from two different mechanisms involving tropical climates. It is important to note that the effects of geography changes over time and historical events can have a disproportionately large effect on certain regions. Areas that were heavily colonized, sub-Saharan Africa and Latin America, had less ethnic plough use and are rather tropical (see map above). Perhaps the colonization process flipped gender roles and create the finding uncovered in this regression. Tropical climate could be inherently tied to fewer women in the labor force. However, colonization changed these roles as colonist gave preferential treatment to females. This later effect is rears itself in increases FLFP in areas that had less ancestral plough use. While this story is entirely speculative (and I don't have the data to test it), it demonstrates that such a mechanism could possibly exist.

Another possibility for this behavior in these geographic variables is that the ethnic geographic control is poorly constructed. While some heterogeneity is expected due to migration, the lack of correlation between present day tropical climate and ethnic tropical climate is concerning. This may result from their tropical climate control not truly representing where people lived, as it was constructed from the fraction of land in tropical climates. In turn it may give some ethnicities a different tropical "score" because and this would bias results. If this was systematically done it could give rise to bias that would contaminate these results.

In either case, present day geography appears to have a positive effect on modern gender roles. By excluding these variables from their controls, the authors could have overlooked any number of mechanism by which geography changes the way women engage with the modern economy. As economies change over time how the environment affects institutions and economic outcomes changes as well. In essence these authors fall into the trap of just accounting for geography at one period in time and not over the whole process of development; overlooking several possible mechanisms that could affect how women engage with the labor force.

It is possible still that both of these controls are skirting around some deeper geographic determinant that is observed through tropical climate. Hence, all that can be concluded from my robustness checks are that the measure of ethnic geography in Alesina et al. is not entirely

precise and that something highly correlated with present day tropical climate has an effect on modern day gender roles that is captured by ancestral plough use.

Conclusion

Alesina, Giuliano and Nunn present several interesting findings concerning the plough and gender roles. They demonstrated that the plough had an effect on pre-industrial labor markets and support Pryor's hypothesis that the plough was an endogenous outcome of geography. The paper falls short in conclusively defining a mechanism by which ancestral plough use affects gender roles. The results from the WVS data and the reduced form instrumental variable regression indicate some omitted variable may be at play. The robustness check shown above show these flaw may be more than noise in the data. These inconsistent findings may be the result of an omitted variable: present day tropical climate or something that is captured by that variable. In conclusion, more work is needed to determine whether plough use affects present day gender roles or if it simply captures some other effect these authors overlooked.

AUTHORS NOTE

I considered rerunning the IPUMS with Ashraf et al. malaria and tropical variables but the bloody data set was too big to compute a merge on Sawyer desktops let alone my old dilapidated laptop. Alas these results could not be checked in the constrained environment. Something to consider going forward.