Social Norms of Work Ethic and Incentives in Organizations

Pedro Forquesato*

Department of Economics
Pontifical Catholic University of Rio de Janeiro.

March, 2015

Abstract

In this paper we model the relation between the dissemination of social norms of work effort (henceforth work ethic) in a given society and the choice of incentives by firms, and we motivate it by presenting evidence from three different datasets that suggests that work ethic is correlated to the intensity of firms' incentives. When the effort choice of different agents in a firm is complementary, having hard-working co-workers makes an agent more productive, and thus in equilibrium he will work harder. Foreseeing this, since work ethic is more useful to hard-working agents, parents will be more willing to transmit it in societies where the probability that their offspring will have co-workers with work ethic is higher. We then expand the model to incorporate firm technology choice, allowing firms to decide between complementary and separable production processes. We show that in societies with high dissemination of work ethic firms will want effort to be complementary (as in modern production processes), while the opposite is true when the dissemination of work ethic is low. Finally, we investigate the comparative dynamics of the model.
JEL Classification: Z10, M50, M52.
Keywords: cultural transmission; work ethic; incentives.

*I thank Vinicius Carrasco, Filipe Campante, Thierry Verdier, Alberto Alesina and Oliver Hart for comments, and participants of seminars at PUC-Rio, the Development Lunch and Research in Contracts and Organizations at Harvard University. Part of this research was done as a visiting scholar at John F. Kennedy School of Government, Harvard University, and I thank them for their great hospitality. That stay was possible thanks to CAPES scholarship, project number 99999.011547/2013-08. Contact e-mail: pedro.forquesato@econ.puc-rio.br.

"Seest thou a man diligent in his business? He shall stand before kings". (Proverbs, 22-29).

1 Introduction

For a long time sociologists and classical economists have defended the importance of differences in cultural values of work effort (henceforth work ethic) as a determinant of differences in development (e.g., Weber (1998)), and a recent literature (Clark (2008), Clark (1987)) has used historical evidence to emphasize the importance of these explanations. There is not in the literature, however, a clear understanding of how these differences in cultural values emerge and disseminate in different cultures.[1]

In this paper we put forward the hypothesis that differences in the dissemination of work ethic emerge endogenously from multiple equilibria in cultural transmission decisions from parents and incentive choice by firms. We build a model that shows how this process would occur, and we substantiate it with anedoctal and empirical evidence that firms' incentives are an important determinant of (intergenerational) cultural choice.

In our model, parents choose to transmit work ethic to their offspring altruistically, taking into account the effect it will have on their children's utility. Where incentives are higher-powered, their children will work harder, and therefore work ethic is more rewarding. In these societies, parents have higher incentive to transmit work ethic, and slowly work ethic disseminates. Incentive steepness has an effect on work ethic.

On the other hand, when effort is complementary among workers, having an employee with work ethic (which in equilibrium will exert higher effort) improves the productivity of his peers. Having more productive workers, the firm optimally chooses to demand more effort from them, for which it must offer higher-powered incentives. If workers' effort are complementary, then in societies in which the proportion of individuals with work ethic is higher, firms choose to offer steeper incentives and demand more effort.

Consequently, we derive two equilibria. When a large proportion of a society's workforce has work ethic, firms want to choose steeper incentives even to those without, since having high effort co-workers makes them more productive. And if firms choose high-powered incentives then agents work hard, and so cultural values that stimulate effort are particularly useful to them. On the other hand, if a small proportion of the workforce has work ethic, then workers shirk more, making their co-workers less productive and causing firms to choose lower-powered

[1]The literature is not unanimous, for example, on the effect of Weber's own explanation for the emergence of these cultural norms. Vide Cantoni (2013) and Becker and Woessmann (2009). We discuss how our explanation reconciles with Weber's in Section 4.

incentives and demand less effort, even from workers with work ethics. Therefore, as individuals work less, work ethic is not as useful, and parents have less incentive to transmit it.

Having established the main argument, we then augment our benchmark model by allowing firms to choose their production technologies, adopting either a complementary technology in which an agent's productivity is highly dependent on his co-workers', or a separable technology, in which interdependencies between workers is minimal. We interpret this choice as between modern production processes, like Fordism or Lean Manufacturing, which are more efficient ways of organizing production, but in which a worker's productivity is highly dependent on his peers' effort, and traditional technologies, like the putting-out system, with less overall efficiency but in which workers are less dependent on their co-workers.

In this scenario the model has two equilibria, one in which firms use traditional technologies and workers do not have work ethic and another where firms adopt modern production processes and work ethic is disseminated in the society. We believe those patterns match stylized historical facts, and that they help explain why countries where labor is unproductive are reluctant to adopt new technologies and organizational structures from developed countries. A similar argument is made in Kremer (1993). Here we contribute by explicitly modeling worker effort exertion and firm incentive choice. That way, our model clarifies also why managers do not employ different incentive schemes in firms where labor productivity is low, and explains the evidence provided in Clark (2008) that it is effort, even more than skill, what differentiates productivity among different regions.

To show the importance of the relation between incentive steepness and work ethic beyond the use of historical examples, we use 3 different data sets and show that higher dissemination of work ethic is correlated with measures of incentive steepness and effort at work. First, using data from the International Social Survey Programme we show that people living in countries where the proportion who believes work is a person's most important activity is high are more likely to arrive from home exhausted, a proxy for effort. We then use data from the European Values Study and European Social Survey to show that in regions where more people believe hard work is an important characteristic to be taught at home, wages are more likely to depend on effort, a measure of incentive steepness. Finally, we go beyond self-reported values by examining the correlation between work ethic and actual data from executive compensation, taken from Standard & Poor's Execucomp, for the United States. The results of these empirical exercises are consistent with our model: workers in areas with higher dissemination of work ethic are more likely to receive high powered incentives and on average work harder, indicating the importance of understanding cultural norms on the behavior of workers and firms.

Our paper is based on a new and growing literature on intergenerational transmission of cultural traits, reviewed in Bisin and Verdier (2011). In particular, in our model parents are imperfectly altruistic in choosing how to socialize their preferences to their offspring, as in Bisin

and Verdier (2000, 2001). Our cultural transmission process, however, has endogenous cultural intolerance, determined by the average incentive scheme chosen by firms, which is a function of the proportion of agents with work ethic in the society. In fact, since parents will want to socialize work ethic when most individuals have it, our model exhibits strategic complementarity in the socialization process (as in, for example, Hauk and Saez-Marti (2002)).[2]

Closely related to work ethic, a growing literature investigates cultural beliefs about the returns of effort, although its focus has been on explaining redistribution choices and social protection (as in Alesina and Angeletos (2005) and Benabou and Tirole (2006)). While these papers provide an important mechanism by which there might be differences in effort and cultural norms of work among different countries, namely political choice of social insurance and redistribution, they don't explain why these differences might exist within countries (as between south and north Italy), where redistribution levels are roughly similar, or in non-democratic societies, where redistribution levels are not decided by majority voting.

Furthermore, our model has different policy implications. If multiple equilibria emerged as a consequence of different levels of redistribution and beliefs on returns to effort, then an exogenous change in taxes and redistribution would cause a change of equilibrium. In our model, however, that is not necessarily true. A reduction of taxes in a region with low dissemination of work ethic might not be able to change the equilibrium, as it might still be optimal for parents not to socialize work ethic to their offspring because firms employ low-powered incentives . In this sense, we contribute by offering an alternative (and complementary) mechanism that can explain differences in effort and that does not depend on political mechanisms. Moreover, the fact that empirical evidence shows that incentive steepness is correlated to measures of work ethic suggests that our mechanism is a relevant one in explaining differences in work behavior between regions.

Perhaps the closest paper to ours is Doepke and Zilibotti (2008). They create a model in which middle-class parents choose to transmit work ethic while the nobility prefers transmitting preferences for leisure. When the Industrial Revolution occurs and work ethics becomes particularly useful, the middle-class takes over and becomes the new entrepreneurs and the dominant economic class. The main difference between their model and ours is that we endogenize the choices of wages and incentives by firms and we focus our analysis on effort exertion, not labor supply. This difference is important both in order to explain the anecdotal evidence that workers' effort is an important determinant of productivity as to explain the empirical evidence that work ethic is highly correlated with workers' incentives steepness, something that can only be meaningfully analysed in imperfect information models.

The rest of the paper is organized as follows. In Section 3 we present empirical evidence that suggests that workers' effort and incentive steepness vary among regions, and that this variation

[2]See Bisin and Verdier (2011) for other references of papers with endogenous cultural intolerance.

is robustly correlated to regional differences in work ethics. In Section 4 we present our model of work ethic and incentives in organizations and argue that it can explain these stylized facts. In Section 5 we discuss some assumptions and results of that model, while in Section 6 we present an important extension of our model, considering the implications of firms being able to choose whether their technology has complementarities, and we analyze the comparative dynamics of the extended model. Finally, we conclude with a discussion of our results and paths for future research.

2 Work Ethic

It is an intuitive stylized fact that countries (and even regions within countries) exhibit very different cultural norms related to work, and sociologists and economic historians have for a long time proclaimed those differences as one of the major explanations for the productivity (and consequently income) differential among nations. Weber, for example, in his book about the protestant work ethic, wrote:

> "As every employer knows, the lack of *coscienziositá* of the labourers of such countries, for instance Italy as compared with Germany, has been, and to a certain extent still is, one of the principal obstacles to their capitalistic development. Capitalism cannot make use of the labour of those who practice the doctrine of undisciplined *liberum arbitrium* ...." (Weber, 1998, p. 21).

Work ethic is understood as a cultural norm that is internalized by agents and gives them a higher cost of shirking, while provides them with a moral support for exerting high effort, easing its cost. Examples of this kind of thinking abound in historical and religious sources. Benjamin Franklin, for example, in his Advice to an Young Tradesman, Written by an Old One, writes:

> "The Sound of your Hammer at Five in the Morning or Nine at Night, heard by a Creditor, makes him easy Six Months longer. But if he sees you at a Billiard Table, or hears your Voice in a Tavern, when you should be at Work, he sends for his Money the next Day." Franklin (2007).

Summarizing the importance of work ethic for capitalistic development, Weber writes: "This peculiar idea (...) of one's duty in a calling, is what is most characteristic of the social ethic of capitalistic culture." (Weber, 1998, p. 19).

Another important source of social norms related to effort is religiosity. Baxter states that "It is for action that God maintaineth us and our activities; work is the moral as well as the natural end of power." (Baxter, 1996, p. 375). But work ethic is not only a social norm that values effort and dedication, but also one that condemns leisure and shirking. Weber, citing Baxter, writes that "Waste of time is thus the first and in principle the deadliest of sins. (...) Loss of time

through sociability, idle talk, luxury, (...) is worthy of absolute moral condemnation." (Weber, 1998, p. 104).

Therefore, while work ethic in an utilitarian sense might be useful for hard working individuals, as it promises ethical, monetary or religious compensation for their toil, on the other hand it might be detrimental to individuals who favour leisure, imposing guilt and moral condemnation for their low effort.

3 The Basic Facts

Although Weber's argument that differences in effort among regions and countries might be substantial is not formal, new empirical evidence seems to substantiate it. Ichino and Maggi (2000), while investigating the behavior of employees of a large Italian bank, discovered that employees of branches in Southern Italy were significantly more likely to shirk then employees of Northern Italy branches. Moreover, they argue that individual preferences between shirking and work explain most of this differential (as compared to sorting, peer effects and firm attributes).

Gregory Clark provides anecdotal evidence that also points to significant differences in worker behavior among countries and regions:[3]

> "In fact workers in poorly performing economies simply supply very little actual labor input on the job. Workers in modern cotton textile factories in India, for example, are actually working for as little as fifteen minutes of each hour they are at the workplace." (Clark, 2008, p. 13)

Furthermore, he argues that this low level of effort inside factories was indeed well known by managers and firm owners. This raises the question of why managers did not incentivize their employees to work harder.[4] By endogeneizing firm choice of incentives (hence, indirectly, worker effort), our model can provide an answer to this puzzle.[5]

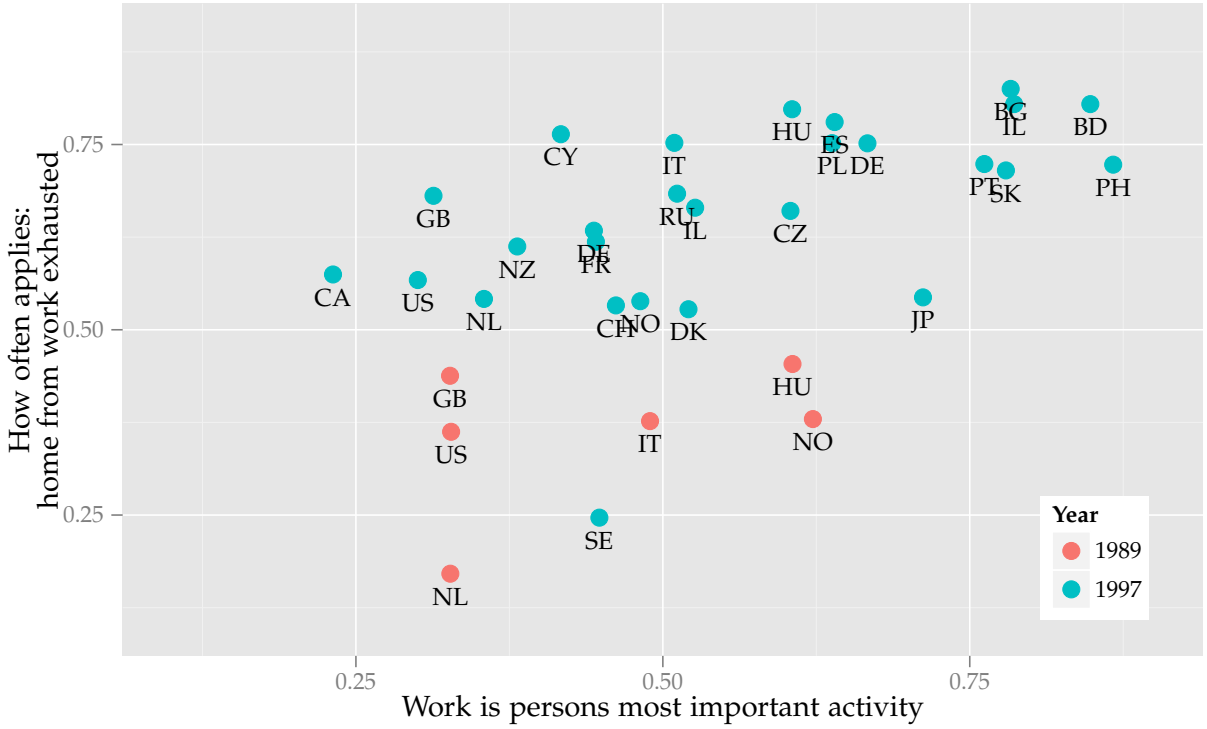
In this section we show quantitative evidence from three different datasets that is consistent with work ethics differing among countries and regions, and in a way correlated both with differences in effort provision and incentives' steepness. We provide descriptive statistics and other information about all our empirical exercises in Appendix B.

[3] For related arguments, see also Clark (1987).

[4] Clark (2008) mentions, for example, that "[t]o partially control this absenteeism some employers used a pass system, under which a worker could leave the mill only with a pass or token from his or her department. Each department was allotted passes equal to 10-25 percent of the staff".

[5] Another possible explanation for this is collusion between managers and workers. Clark offers evidence that this mechanism shouldn't have been very important: for example, in 1895 in Bombay, of 55 mill managers, 27 were British, as were 77 of the 190 weaving masters and other managerial positions. It seems implausible that British managers would be more willing to collude with Indian workers than British ones.

Figure 1: Cross-country correlation between work ethic and effort at work (ISSP 1989, 1997)



3.1 International Social Survey Programme

First, we use International Social Survey Programme: Work Orientations I and II (ISSP 1989 and ISSP 1997) data on work ethic and effort.[6] The International Social Survey Programme is a cross-national collaboration on social science surveys. We use it to build individual-level and country-level measures of work ethic and explore the correlation of these variables with a proxy of work effort by individuals.

As a measure of work ethic we analyse respondents' answer to the question: "How much do you agree with or disagree with, thinking about work in general: Work is a person's most important activity?" For more easily interpretable results, we create a dummy variable encoding one if the respondent answers "Strongly agree" or "Agree" and zero otherwise. As a measure of work effort, we use the respondent's answer to the question "How often do you come home from work exhausted?" If the respondent answers "Always" or "Often" we give the variable value one, and zero if he replies "Sometimes", "Hardly Ever" or "Never."[7]

The correlation between those two variables is displayed in Figure 1. It shows that both work

[6] We don't use the 2005 wave because the questionnaire does not contain the questions we are interested in.
[7] We get the same qualitative results using the original variables.

ethic and effort vary significantly among countries in both years, although work ethic seems to vary little from 1989 to 1997 in countries with both values, which is consistent with our view of culture as slow to change. Moreover, there is a correlation between how important people in a country consider work to be and how much effort people on average exert on their work, and this correlation seems to hold for both years we have data for. We provide more careful evidence in this direction below.

As in our model, we consider the relevant dimension of work ethic to be its average level in the population, since this affects an individual's probability of having co-workers with work ethic. We therefore aggregate the measure of work ethic by country. We also use a number of individual and country-level controls. Table 1 uses individual-level data to formalize the argument of Figure 1. It shows that the correlation is statistically significant and robust to the addition of controls.

The first column shows that individuals with work ethic are significantly more likely to return home exhausted from work, which is an indicator of how much effort they exert on their jobs. On columns two and three we show that this result is robust to addition of individual-level controls and country dummies. Although the point estimates are smaller in both scenarios, they are still statistically significant at the 1% level.

Table 1: Correlation between work ethic and effort at work (ISSP 1989, 1997)

	(1)	(2)	(3)	(4)	(5)	(6)
Work important	0.0883	0.0432	0.0167	0.0522	0.0163	0.0128
	(0.0118)***	(0.0126)***	(0.0071)*	(0.0068)***	(0.0062)**	(0.0062)*
Work imp. (country)				0.3480	0.3067	0.5739
				(0.0861)***	(0.1068)**	(0.0600)***
Education (yrs)		-0.0138	-0.0110		-0.0107	-0.0110
		(0.0030)***	(0.0024)***		(0.0028)***	(0.0024)***
Female		0.1047	0.1085		0.1075	0.1091
		(0.0088)***	(0.0083)***		(0.0081)***	(0.0082)***
Year (1997)	0.2762	0.2628	0.2530	0.2578	0.2247	0.2469
	(0.0290)***	(0.0357)***	(0.0478)***	(0.0281)***	(0.0388)***	(0.0295)***
Individual-level controls	N	Y	Y	N	Y	Y
Country dummies	N	N	Y	N	N	Y
R^2	0.058	0.124	0.163	0.069	0.132	0.170
N	40617	32018	39694	40617	32018	32018

Dependent variable is the respondent's answer to: "How often do you come home from work exhausted?", where we recode "Always" or "Often" as one, and "Sometimes", "Hardly Ever" or "Never" as zero. Standard errors are corrected for cluster at the country level. P-values are significant at: 0.05 (*), 0.01 (**) and 0.001 (***) levels. Individual-level controls are years of education, sex, age, age squared and respondent's self-reported social class. There are 31 country-level observations.

On columns (4)-(6) we analyze the effect of country-level work ethic on workers' effort choice.

As our model predicts, individuals with work ethic exert more effort, but the environmental culture is also important (in fact our regression suggests it is even more important) in determining their effort. Column five uses individual-level controls, and column six both individual-level and country fixed-effects.[8] The work ethic effect changes little in (5), and actually increases significantly after we control for country fixed-effects.

A clear *caveat* of this analysis is that both measures of work ethic and effort are self-reported. It is possible, therefore, that the results are driven by endogenous measurement errors. If we believe those measurement errors are random, however, then we would expect attenuation bias, which works against our predictions. The negative coefficient for years of education and strong positive coefficient for women points to the flaws of our measure of effort. In the first case, a possible explanation would be that more educated workers have more intellectual jobs, which are also less tiresome. In the second case, while possibly women exert more effort than men, an alternative explanation would be that they are more easily exhausted, and it is impossible to discern both effects. While we cannot know if people with work ethic are sorted into more tiresome jobs than those without (for some reason other than their willingness to exert higher effort), as long as we believe that those with work ethic are not more prone to feeling exhausted (or reporting it) than workers without work ethic, our measures will be underestimating the total importance of work ethic on effort.

3.2 European Social Survey and European Values Study

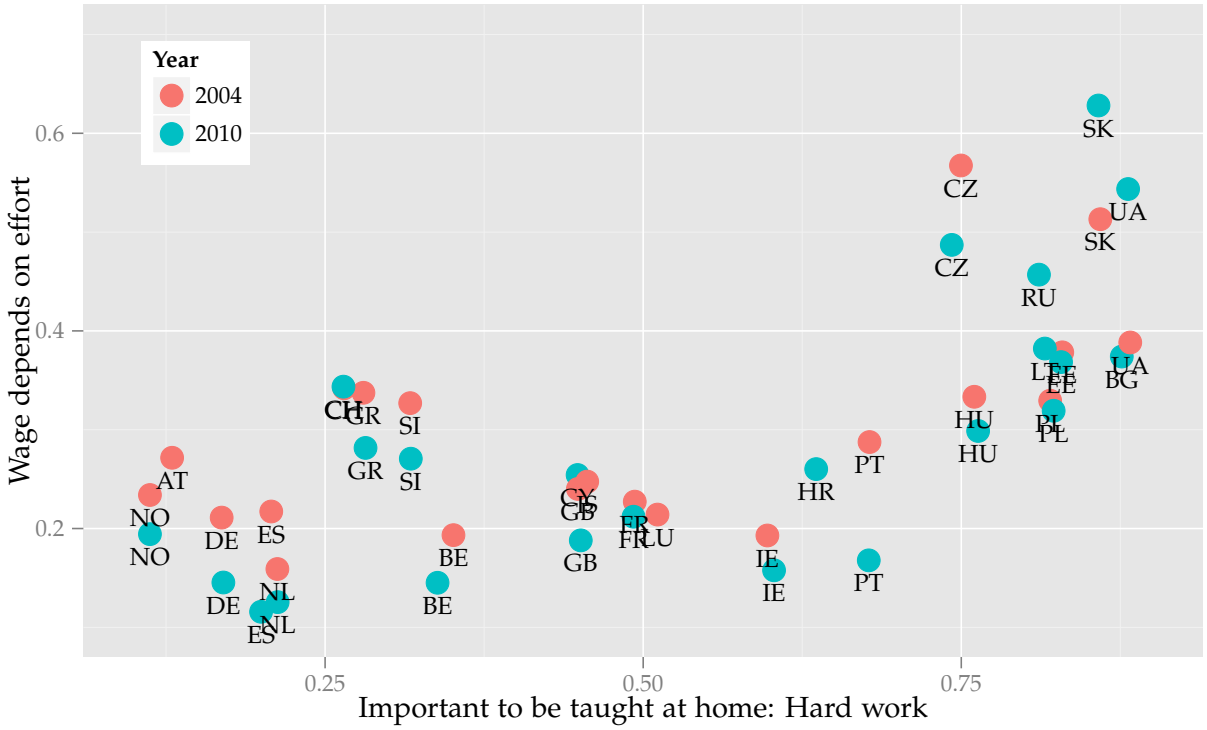
Our second empirical exercise uses data from European Social Survey: Family, Work and Well Being surveys (ESS2 2004 and ESS5 2010) and European Values Study 4th Wave (EVS 2008) to form a database on regional level work ethic variables and individual-level work attitudes. The advantage of this database is that the ESS has questions which are plausible indicators for steepness of incentives (we are interested in how high-powered are the incentives, not how large are the wages, something more easily measurable).

From the EVS 2008 we aggregate at the regional level an indicator of work values, namely if the respondent mentioned "Hard work" when asked, "Here is a list of qualities which children can be encouraged to learn at home. Which, if any, do you consider to be especially important? Please choose up to five!" We consider that this variable best captures the intergenerational aspect of cultural transmission of work ethic, nonetheless our result is robust to using other questions which are related to these values.

We use the ESS (2004, 2010) data to get individual self-reported measures of steepness of incentives. We use respondents' answer to the question: "Please tell me how true each of the following statements is about your current job. My wage or salary depends on the amount of

[8]Since we have only 31 country-level observations, this limits our ability to control for time-varying country-level variables.

Figure 2: Cross-country correlation between work ethic and incentives steepness (EVS 2008, ESS 2004, 2010)



effort I put into my work." To ease interpretation, again we encode the variable as dummies.

Figure 2 below shows the correlation between self-reported incentive steepness and country-level work ethic for European countries in 2004 and 2010. The relationship seems strong, as the probability the respondent considers his wage highly dependent on effort in countries with high work ethic is twice as big as in countries with low work ethic.

We test more rigorously this correlation in Table 2, using regional-level data on work ethic and individual-level data on self-reported incentives from EVS and ESS, showing that this correlation is strong and robust to a multitude of controls. In our most conservative specification, moving from the 1st quantile to the 3rd quantile in work ethic makes an individual 25 percent (7 p.p.) more likely to think his wage depends on his effort.

Since Figure 2 suggests Eastern European (former Soviet) societies and Western Europe countries have different views regarding incentive steepness and work ethic, in all our regressions we control for belonging to East Europe (our results are only stronger otherwise).[9] Importantly, the

[9]While not central to our argument, we conjecture that the higher prevalence of work ethic among Eastern European countries results from cultural promotion of work values by the socialist government. It is also noteworthy that incentive wages (piece-rates) were prevalent in the industry during the Soviet Union.

regressions show that while this difference is indeed significant, our results are not driven by that.

Unfortunately, work ethic does not vary enough within countries for us to obtain significant results with country dummies. While we believe this is due to culture being more widespread than our local measure of regions and countries in Europe being too small to allow for big variations within, this is also consistent with the view that other country-level variables are causing both higher work ethic and more high-powered incentives. To try to confront these fears (in a limited way), in columns (3) and (4) we control for a variety of region and country-level covariates, including other cultural traits that might influence both the dissemination of work ethic and incentives. While many of these covariates have significant effects, they change little our estimates.

3.3 World Values Survey and Execucomp

Finally, in our third empirical exercise we test the relationship between work ethic and the steepness of incentives provided by firms without using self-reported values. To do so, we use Standard and Poors' Execucomp database on executive compensation for the United States of America. We match that database with USA regional data on World Values Survey (WVS). Execucomp has executive compensation data since 1994 for all S&P 1500 firms, from which we use the detailed executive compensation data to create measures of incentive steepness. Obtaining reliable measures of incentive steepness is difficult for the majority of workers, since their contracts usually involve efficiency wages and career concerns. Data for executive compensation, however, is widely available, discriminated by incentive type, and, since most executives receive some sort of bonuses, more reliable.

We use WVS waves of 1995, 1999 and 2006. For each wave, we match it to executive data from the year before the survey, the year of the survey and the year after the survey (e.g. for the first wave we use 1994, 1995 and 1996).[10] A difficulty with this dataset is that we only have data on work ethic disaggregated by region.[11] Still, since we have data for all three waves and nine years of executive compensation it is possible to get precise estimates, even controlling for regional fixed effects.

On columns (1), (2), (4), and (6) we use as a measure of incentive steepness bonuses over bonuses and wages. It is interesting to note that increasing controls increases the point estimates, suggesting that simple correlations are downward biased. On columns (3) and (5) we take into account that good executives might try to protect themselves from risk and we consider also options and other incentives packages. Our measure of incentive steepness is then $1 - \frac{\text{wage}}{\text{total payment}}$.

[10] The results are robust to using only executive data for the year of the wave and using only some of the waves.

[11] There are ten regions in the data: East South Central, West South Central, New England, South Atlantic, East North Central, West North Central, Rocky Mountain State, Middle Atlantic States, Northwest and California.

Table 2: European Values Study and European Social Survery (EVS 2008, ESS 2004, 2010)

	(1)	(2)	(3)	(4)
Work imp. (region)	0.1304	0.1430	0.1473	0.1143
	(0.0285)***	(0.0295)***	(0.0349)***	(0.0327)***
Female		-0.0584	-0.0596	-0.0603
		(0.0073)***	(0.0072)***	(0.0072)***
Education (yrs)		0.0002	-0.0009	-0.0007
		(0.0014)	(0.0013)	(0.0013)
Self-employed		0.6070	0.6000	0.5937
		(0.0428)***	(0.0546)***	(0.0637)***
Left-right (region)			-0.0522	-0.0522
			(0.0157)***	(0.0140)***
Bribe just (region)			0.0660	0.0722
			(0.0547)	(0.0524)
GDP pc (region)			-0.0162	-0.0002
			(0.0128)**	(0.0143)
Education (region)			0.0618	0.0447
			(0.0162)***	(0.0157)**
GDP pc (country)				-0.0549
				(0.0314)**
Industry % (country)				0.0022
				(0.0012)
East-west (east)	0.1217	0.1138	0.1242	0.1138
	(0.0174)***	(0.0199)***	(0.0227)***	(0.0224)***
Year (2010)	-0.0286	-0.0214	-0.0240	-0.0313
	(0.0087)*	(0.0125)	(0.0118)*	(0.0120)**
Country-level controls	N	N	N	Y
Region-level controls	N	N	Y	Y
Individual-level Controls	N	Y	Y	Y
R^2	0.040	0.085	0.093	0.095
N	34375	16677	16664	16664

Dependent variable is the respondent's answer to: "Please tell me how true each of the following statements is about your current job. My wage or salary depends on the amount of effort I put into my work", where we encode one for "Very True" or "Quite True" and zero otherwise. Standard errors are corrected for cluster at the regional level. P-values are significant at: 0.05 (*), 0.01 (**) and 0.001 (***) levels. Individual-level controls are religion, gender, age, age squared, years of education, employment relation, NACER2 level industry controls and whether there is a child home. Region-level controls are (log) income per capita (PPP), average years of education, average political tendency (left or right), response to whether bribes are sometimes justified, average concern for others, which is more important: freedom or equality and view of the government. Country-level controls are (log) GDP per capita and industry participation. There are 30 countries in the sample.

Table 3: WVS for USA and Execucomp (WVS 1995, 1999, 2005, Execucomp 1994-2006)

	(1)	(2)	(3)	(4)	(5)	(6)
Work imp. (region)	0.1386	0.2156	0.1620	0.1884	0.0982	0.0975
	(0.0584)*	(0.0912)*	(0.0747)*	(0.0625)**	(0.0622)	(0.0445)*
Age		0.0082	0.0201	0.0084	0.0200	0.1609
		(0.0017)***	(0.0014)***	(0.0016)***	(0.0015)***	(0.1596)
Age squared		-0.0001	-0.0002	-0.0001	-0.0002	-0.0015
		(0.0000)***	(0.0000)***	(0.0000)***	(0.0000)***	(0.0015)
Interlock		-0.0507	-0.0671	-0.0507	-0.0687	
		(0.0206)*	(0.0198)***	(0.0214)*	(0.0030)***	
CEO (no)		-0.0326	-0.0641	-0.0329	-0.0647	0.1241
		(0.0025)***	(0.0032)***	(0.0026)***	(0.0035)†	(0.2597)
Year dummies	Y	Y	Y	Y	Y	Y
Region FE	N	N	N	Y	Y	N
Executive Controls	N	Y	Y	Y	Y	*
R^2	0.189	0.345	0.181	0.350	0.186	0.819
N	60518	29492	29319	29492	29319	30

Dependent variable in columns (1), (2), (4) and (6) is $\frac{\text{BONUS}}{\text{BONUS+SALARY}}$. For columns (3) and (5) dependent variable is $1 - \frac{\text{SALARY}}{\text{TOTAL WAGE}}$ (the difference among them is the inclusion of options and other compensation methods). P-values are significant at: 0.05 (*), 0.01 (**) and 0.001 (***) levels. For each wave t of the WVS we use executive compensation data for the years $\{t-1,\ t,\ t+1\}$. Executive-level controls are age, age squared, gender, if she is CEO, if she is CFO, if the board is interlocked and detailed SIC industry controls. All standard errors are clustered at the regional level. Last column aggregates the variables by region and WVS year. Given the extremely low number of observations we chose to remove some controls (namely gender and interlock), besides for obvious reasons we remove the SIC controls. There are 10 regions in the sample.

The results are weaker, but still consistent with our hypothesis. In our preferred specification, (4), an increase from first to third quartiles in regional work ethic increases in 1.5 p.p. the proportion of bonuses in the compensation contract, a five percent increase over the mean.

We use very specific industry-level controls in order to face the possibility that we are capturing different compensation behavior from different industries that are spatially sorted. Furthermore, not only our analysis is within a country and so our results are controlled for country-level institutional features, but since we have panel data we can also control for regional fixed effects. We believe this can account for many of the worries originating from our previous empirical exercises. Columns (4) and (5) show that using region fixed-effects has small effect over our estimates.

Finally, having so few regions may induce doubts about the robustness of our results. In column (6) we apply the most challenging test possible, we aggregate the observations by WVS wave and region, obtaining 30 observations (three waves for each ten regions). Surprisingly, the correlation still holds, although we have to drop the variables "interlock" and "gender" to

maintain power.[12] We believe this supports the robustness of our correlations.

Naturally, using executive compensation raises the question of how the results generalize to the rest of the working population. We hope that future research will use better data and be able to test more carefully the implications of our model, focusing particularly on causal identification and external validity.

4 The Model

In our model there is a population of measure two consisting of overlapping generations of agents that live two periods. In the first period they acquire by a cultural transmission process a binary trait $k \in \{W, N\}$ that signifies the possession of work ethic (W) or not (N). In the second period they are employed by a firm, where they work and earn wages, and also choose how to educate their offspring in work ethic. For simplicity we assume that reproduction is asexual and each individual has exactly one child. Therefore the measure of adults in the population is one and we denote the proportion of adults with work ethic as x^W.

We assume work-related decisions and the cultural transmission decision are separable, and consider then separately. We start by analyzing the effort choice of the agent.

4.1 The Moral-hazard-in-teams Game

We model the individuals' employment in firms as a moral-hazard-in-teams game. There is a continuum of measure one of firms, with each firm being matched with two randomly chosen individuals from the population each period in order to engage in production.[13] We assume that firms employ agents in a project with value ρ and probability of success given by $\theta a_i^\alpha a_j^\alpha$, with $\alpha \in (0,\ 1)$, $a_i \in [0,\ 1]$ and $\theta \in (0,\ 1]$, where the first condition guarantees concavity of the firm's decision problem and the second and third conditions are sufficient to ensure $\Pr(\text{success}) \leq 1$. In this equation a_i (respectively a_j) represents the effort exerted by agent i (resp. j), and θ represents the efficiency of effort of agents i and j. With abuse of notation we'll use the subscripts to refer both to an agent and to his type (trait).

The functional form chosen entails that the probability of success is supermodular, representing a production process in which the effort of workers is complementary. An extreme example where this might be relevant would be a Fordist production plant, where the productivity of a worker in the production line is directly related to the speed in which he receives production inputs, which are the outputs of the previous workers on that line.

[12] These variables suffer from being very homogeneous. Only 5% of executives in our sample are female, and 1.7% of boards are interlocked. We also drop SIC Industry controls, for obvious reasons.

[13] The fact that the measure of firms and workers is the same, yet each firm is matched with two agents brings no mathematical difficulties.

An employee's effort is non-observable, and so the firm must choose a compensation scheme that is incentive compatible. We assume that the firm has complete information on whether the agent has work ethic or not.[14] Since the project outcome is binary, the firm can restrict attention to linear incentive schemes $\{w_{ij},\ b_{ij}\}_{i,j\in\{W,\ N\}}$. Both the workers and the firm are risk neutral.

Agents have a quadratic cost of effort that depends on the trait they acquired in their childhood. Those with work ethic have lower marginal cost of effort $\kappa_W a_i$ as compared to workers without ethic ($\kappa_N a_i$), $\kappa_W < \kappa_N$, representing moral justification of effort that eases their toil. On the other hand, we assume this cannot be dissociated from a moral condemnation towards leisure. For example, in the religious case, while the promise of eternal recompense for hard labour (as expressed in our epigraph) can ease the costs of effort, it cannot be dissociated from the ascetic view that leisure is worthy of "absolute moral condemnation", as Baxter claims. To account for this, we assume that agents with work ethic suffer a cost K_W that is external to their work behavior, indicating an inability to enjoy free time, family and friends.

Thus the utility of an agent of type i with a co-worker type j (defining $\mathbb{I}_A$ as the indicator function of A) is:

$$\left(\theta a_i^\alpha a_j^\alpha\right) b_{ij} + w_{ij} - \mathbb{I}_{\{i|i=W\}} K_W - \kappa_i \frac{a_i^2}{2}. \tag{1}$$

We further assume that the firm can only choose positive base wages and bonuses (i.e. limited liability). Without this assumption, firms would choose negative lump-sums to equalize every employee's payoff to his outside option. By assuming they are legally or institutionally constrained from doing so, we effectively allow employees to capture part of the surplus. This assumption is essential for our results, as otherwise an agent's payoff would be independent of whether he has work ethic or not (as it would always match his outside option).

Moreover, throughout the paper we assume that the outside option is redundant given the limited liability (i.e. the outside option $\overline{u} \leq 0$). Adding a binding outside option (as long as it is not high enough to make the model trivial) would not change the results. These assumptions ensure that $w_{ij} = 0$ for all $i,\ j$.

The firm will choose how much effort to incentivize its employees to exert, knowing that it must compensate them in an incentive compatible way (and respecting limited liability and individual rationality constraints). Defining $b_{ij}(a_i,\ a_j)$ as the bonus necessary to induce action a_i from a worker with co-worker exerting effort a_j, the firm solves:

$$\max_{a_i, a_j} \theta a_i^\alpha a_j^\alpha \left(\rho - b_{ij}(a_i,\ a_j) - b_{ji}(a_j,\ a_i)\right) - w_{ij} - w_{ji}. \tag{2}$$

Subject to:

$$a_i \in \arg\max_a U(\kappa_i, \kappa_j) \equiv \left(\theta a^\alpha a_j^\alpha\right) b_{ij} + w_{ij} - \mathbb{I}_{\{i|i=W\}} K_W - \kappa_i \frac{a^2}{2},$$

[14] We discuss this and other main assumptions of our model in the next section.

$$a_j \in \arg\max_a U(\kappa_j, \kappa_i) \equiv (\theta a^\alpha a_i^\alpha)\, b_{ji} + w_{ji} - \mathbb{I}_{\{j|j=W\}} K_W - \kappa_j \frac{a^2}{2}, \text{ and that}$$
$$w_{ij} \geq 0,\ w_{ji} \geq 0,\ U(\kappa_i, \kappa_j) \geq \overline{u}, \text{ and } U(\kappa_j, \kappa_i) \geq \overline{u}.$$

From the incentive compatibility constraint it is straightforward to calculate the bonus necessary to implement a given effort level. As the worker only receives the bonus when the project is successful, we define μ as the expected compensation received by the agent. We then have that:[15]

$$b_{ij}(a_{ij},\ a_{ji}) = \frac{\kappa_i a_{ij}^2}{\theta\alpha a_{ij}^\alpha a_{ji}^\alpha} \text{ and } \mu(a_i) = \frac{\kappa_i a_{ij}^2}{\alpha}. \tag{3}$$

The technical tractability of this functional form comes from the expected compensation μ received by an agent not depending (directly) on his co-worker's effort.

Since there are complementarities in effort, when an agent has a hard-working co-worker, his own effort becomes more productive, and the firm will choose to incentivize him to work harder as well. In fact, maximizing its profit, the firm will choose to request from agent i with co-worker j effort:

$$a_{ij}^* = \left[\frac{\theta\rho\alpha^2 (a_{ji}^*)^\alpha}{2\kappa_i}\right]^A = \left[\frac{\left(\theta\rho\alpha^2\right)^{1+\alpha A}}{2^{1+\alpha A}\kappa_i \kappa_j^{\alpha A}}\right]^{\frac{A}{1-\alpha^2 A^2}}. \tag{4}$$

Where $A = \frac{1}{2-\alpha}$. Intuitively, a higher cost of effort for agent i makes him more costly to incentivize, and the firm will choose to propose less effort. On the other hand, higher efficiency θ and higher project value ρ make the agent more productive, and the firm will propose higher effort. More important to our argument, the agent's co-worker's effort has direct effect on the agent's productivity, hence having a high effort co-worker causes the firm to propose higher effort to the agent, and by equation (3) his total compensation will increase.

As the only form of heterogeneity is in cost of effort, which represents the agents' work ethic, the optimal contract is entirely determined by the agent's work ethic and his peer's. This is represented in the second part of equation (4). Workers without work ethic and workers with peers without work ethic in equilibrium work less. Even more interestingly, equation (4) shows that while all individuals work harder when their peer has work ethic, this difference is greater for workers with work ethic themselves. Furthermore, equation (3) shows us that the same is true for the total compensation received by workers: individuals matched with a peer with work ethic earn higher compensation than those matched with a peer without work ethic, but this effect is greater for individuals with work ethic themselves.

Intuitively, in a Fordist production line the faster that inputs reach a worker station the faster

[15] Although the bonus paid to the individual is decreasing in his co-worker's effort, the fact that his own proposed effort is higher when the co-worker has work ethic (see equation (4)) makes the bonus increasing in co-worker's work ethic. Therefore, regions with high dissemination of work ethic indeed face higher-powered incentives. We discuss this with more detail in Section 5.

that the employee can work. But this is only useful inasmuch as she can produce at least as fast as the inputs arrive. While for fast workers the advantage of being in a fast production line (when earning piece rates) is very large, for a slow worker it is negligible.

We now proceed to the main result of this section, which is a formalization of the discussion above. For that, we define the *matching payoff* of a worker type i when matched with co-worker type j ($i,\ j \in \{W,\ N\}$), and we denote it by $V(i,\ j)$, the utility the worker type i receives under the optimal compensation scheme when matched with a worker with type j.

Using the matching payoff, we can define the *ex-ante utility* of an individual of being of type i ($i \in \{W,\ N\}$) as being:

$$U^i(x^W) = x^W V(i,\ W) + (1 - x^W)V(i,\ N), \tag{5}$$

where x^W represents the proportion of adults with work ethic in the society. The ex-ante utility is the expected utility an individual will receive in the labor market, given his type, but without knowledge of whether his co-worker will have work ethic or not.

The main result in this section is stated below and proven in Appendix A (as well as all other Theorems in the main text).

Proposition 4.1. The benefit of having work ethic is greater for workers whose peers also have work ethic. Formally, the matching payoff function is increasing differences in types, i.e. $V(W,\ W) - V(N,\ W) > V(W,\ N) - V(N,\ N)$.

4.2 Cultural Transmission Process

Having established the labor market payoff of different cultural traits, we can define the problem of the parent choosing whether to transmit his cultural trait to his offspring. As work ethic is an internalized social norm that affects directly the preferences of individuals, we follow the literature established by Cavalli-Sforza (1981), Boyd and Richerson (1985) and Bisin and Verdier (2000, 2001).

Parents of type i choose a socialization effort $d^i \in [0,\ 1]$ in order to attempt to pass their trait on to their offspring. This socialization effort succeeds with probability d^i, in which case the offspring will acquire the parental trait (vertical socialization). If this effort is unsuccessful, the child will acquire the trait from a random element of the population (horizontal/oblique socialization). In that case, the child will acquire the parent's trait with probability x^i, where x^i is the proportion of the population with trait i. Consequently, the probability the child of a parent type i acquires trait i (resp. j) is given by:

$$P^{ii}(d^i;\ x^i) = d^i + (1 - d^i)x^i$$

$$P^{ij}(d^i;\ x^i) = (1 - d^i)(1 - x^i)$$

Parents are altruistic and consider their offspring's expected utility when deciding whether to transmit work ethic. As we argued above, the ex-ante utility of having work ethic is a function of the proportion of individuals with work ethic in the population. We assume, however, that parents are myopic: they choose the socialization effort based on the dissemination of work ethic they observe in the society. This can be seen as a form of imperfect altruism, in the sense that parents observe their own environment and payoff when considering their children's utility.

Parents of type i therefore maximize with regards to socialization effort d^i:

$$P^{ii}(d^i;\ x^W)U^i(x^W) + P^{ij}(d^i;\ x^W)U^j(x^W) - \frac{(d^i)^2}{2}$$

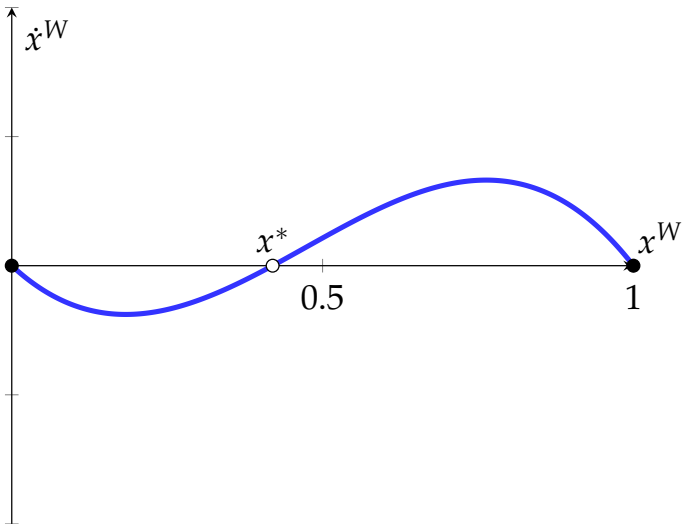
where U^i is the ex-ante utility of individual of type i introduced in last section. First order conditions imply that $d^W = (1 - x^W)\Delta U(x^W)$. As in the literature, we call $\Delta U(x^W) \equiv U^W(x^W) - U^N(x^W)$ the cultural intolerance of work ethic. But as it is clear by equation (5), the cultural intolerance is endogenous; it depends on the proportion of the population with work ethics. Moreover, Proposition 4.1 shows that individuals with work ethic gain more by having a peer with work ethic than individuals without work ethics. This is therefore a game of strategic complements, and our cultural transmission process is a process of cultural conformity (see Bisin and Verdier (2011)). We formalize this assertion in Proposition 4.2 below.

Proposition 4.2. Assume parameters are not such that it is always better or always worse for an agent to have work ethic, independently of whom he works with. Then there exists $x^* \in (0,\ 1)$ such that if the dissemination of work ethic in a moment t, x_t^W, is less than x^* the proportion of individuals with work ethic converges to zero, and if x_t^W is larger than x^*, it converges to one.

Figure 3 below illustrates Proposition 4.2 for a set of parameters. In societies with low dissemination of work ethic, parents expect their children to be more likely employed with low effort co-workers and paid low-powered wages. Doing so, they find it in their offspring's best interest to transmit cultural traits opposed to work ethic, that promote leisure and the importance of personal life as opposed to the career. On the other hand, in societies with high dissemination of work ethic, the average compensation scheme is high-powered, and so the returns for high effort are large. In this case, parents will altruistically choose to disseminate work ethic, propagating beliefs that work is important and that it provides a purpose in life (a *calling*).

Figure 3 also highlights two of the main features of our model: that history matters, in the sense that the dynamic equilibria are path dependent, and that very similar initial conditions can lead to drastically different equilibrium cultural outcomes. Together, these characteristics imply that cultural and economic shocks in the past can have - through cultural propagation - effect on current economic outcomes. In particular, it explains what Weber (1998) called the "ghost of dead religious beliefs", the hypothesis that Protestantism marked the dissemination of work ethic among Northern European societies, but that this work ethic continued to expand and propagate

Figure 3: Dynamics of work ethic for an example parameter profile. ($\alpha = 0.7$, $\theta = 0.25$, $\rho = 10$, $\kappa_N = 0.8$, $\kappa_W = 0.4$, $K_W = 0.6$)



even after lost importance in Western religion the view of work as a means to salvation. Although initially work ethic disseminated through a religious shock, as individuals became more hard-working, firms became more productive and started employing steeper incentives, causing the transmission of work ethic to be beneficial even after it lost its religious meaning.

Furthermore, as we argue in Section 6.1, parental socialization and firm technological choice are related: after the Protestant Revolution, firms chose production processes and managerial techniques focused in complementarities among workers to increase the productivity effect of a now more hard-working population, which then made the dissemination of a culture that emphasize hard work more valuable.

In this Section we provided a model which predicts multiple equilibria in effort, incentives and dissemination of work values in the society. This model is consistent with empirical findings that work ethic is correlated with measures of effort and incentive steepness, with previous literature that emphasized the importance of personal background to explain differences in work effort among regions, and with stylized facts in the literature that point to significant importance of effort exertion in explaining productivity differential among nations and different periods. In the next section we discuss some assumptions of our model, and in Section 6 we extend this model to include firm technological choice.

5 Discussion

In this section we discuss some assumptions and features of our model. First, we argue that the assumption that the firm knows the types of the workers and can offer different wages for different types is reasonable and matches available empirical facts. Then, we argue that the simplification of our model by considering only bonuses implies no loss of generality, even though only a small fraction of the work force actually receives incentive wages. We also discuss the essential assumption of complementarity in effort and finally we discuss welfare consequences of our model.

Imagine that the firm can observe the output y of the agent each period, but not the effort. If the effort is constant, as in our model, then by observing the output a sufficient number of times the firm can infer, to any degree of chance, the type of the agent. Therefore, if the contractual relation is sufficiently long, it is with no loss of generality to suppose that the firm has knowledge of the type of the agent, even if at any given point in time it cannot observe the effort the agent supplies.

Moreover, if when the output is realized the firm not only pay bonuses, but also adjusts their rate (for example, by raises), then by choosing an appropriate adjustment rule with time she can make the payment of different types of agents converge to their optimal values, even if direct discrimination is somehow forbidden or discouraged. Consequently, there is no loss of generality in assuming that firms can freely choose different optimal bonus levels for different combinations of workers' type.

Another important element of our model is that the firm pays bonuses to workers to incentivize them to exert more effort, while in reality most workers do not receive any kind of incentive wage. This is consequence of the fact that in reality many other types of incentives are available, most commonly career concerns and efficiency wage models. While our model does not explicitly account for these other forms of incentives, the intuition is nonetheless the same, and the generalization would be straightforward. Therefore, while we only directly model piece-rates workers, our argument is general.

Important for the interpretation of our model is to note that substituting equation (4) in equation (3) we obtain:

$$b_{ij}(\kappa_i,\ \kappa_j) = \frac{\kappa_i}{\theta\alpha}\left[\frac{\left(\theta\rho\alpha^2\right)^{1+\alpha A}}{2^{1+\alpha A}\kappa_i\kappa_j^{\alpha A}}\right]^{(2-\alpha)\frac{A}{1-\alpha^2A^2}}\left[\frac{\left(\theta\rho\alpha^2\right)^{1+\alpha A}}{2^{1+\alpha A}\kappa_i\kappa_j^{\alpha A}}\right]^{-\alpha\frac{A}{1-\alpha^2A^2}} \tag{6}$$

Regarding the effect of having a co-worker with work ethic, there are two opposite forces at work here. On the one hand, since the marginal cost of effort is lower, for a given level of effort the firm needs to compensate the worker less for it, and so the bonus will be smaller. On the other hand, the firm will want to incentivize the worker to exert more effort, and will need to

pay him a higher bonus for that. With some algebra, equation (6) shows that the second effect is stronger: a lower cost of effort for a worker and for his co-worker will result in higher bonuses, that is, higher powered incentives. In this sense our model predicts that societies with higher dissemination of work ethic have higher powered incentives.

A further assumption of our model is that effort is complementary among workers. Complementarity is a topic that receives increasing attention in the literature of organizational economics (e.g. Milgrom and Roberts (1990), Bresnahan et al. (2002) , Brynjolfsson and Milgrom (2012)). Although observational evidence points to complementarity between agents' effort as being an important part of modern production processes (for example, Fordism and Lean Manufacturing), complementarity in effort has not been an important part of models of organizational behavior so far.

One reason is that adding complementarities results in significant loss of tractability in models of moral hazard, and in this sense another contribution of this paper is to suggest a tractable model of effort complementarity. A second reason is the difficulty in obtaining causal data on the existence of complementarity effects: although observational evidence points to sorting on worker productivity and diligence among sectors and firms, it is difficult to disentangle the effect coming from complementarities between workers from other forms of complementarities.

Adding complementarities to our model has one intuitive and attractive implication, however. In our model an agent's wage depends on the effort chosen by his co-workers, and this dependence is higher for hard working agents. This is consistent with wages being different between firms and sectors, and correlated inside a firm, features that we observe empirically. Groshen (1991), for example, notes that occupation and employer identity explain 90 percent of variation in wages.

Our model has well-defined implications for welfare, even though welfare comparisons in models with changing preferences are notoriously problematic (see Stigler and Becker (1977)). According to Proposition 4.2, there are only two possible asymptotically stable dynamic equilibria: one with full dissemination of work ethic and another with no work ethic in the society; welfare analysis is reduced to comparing these two equilibria.

But indeed in our model the dynamic equilibria (of Proposition 4.2) are ranked in terms of welfare, as all workers (and firms) prefer to have co-workers with work ethic. Even when having a hard-working co-worker does not affect an agent's payment directly, as is the case with the functional form we analyze in this paper, it does make the agent more productive, and more productive agents work harder, receive a larger part of the surplus, and have higher utility. Our model, therefore, represents a world in which effort pays; individuals who work harder have higher utility.[16] Hence, for any parameter profile for which it is not always worse to have work ethic, the equilibrium with full dissemination of work ethic is more efficient, as by the argument

[16]It is important to note this is in fact a general property of moral hazard problems with limited liability. To see

above and Proposition 4.1: $V(W,\ W) > V(N,\ W) > V(N,\ N)$.

6 Extensions

In this section we analyze possible extensions to our benchmark model. A question raised by our previous analysis is why firms in societies with low dissemination of work ethic would choose complementary technologies, as this would cause a drop in productivity. In Section 6.1 we augment our base model to account for firm technology choice and we show that indeed firms in low work ethic societies will prefer to have non-complementary technologies. We then characterize the phase diagram induced by the firm technological choice and parental socialization transmission. Finally, if we interpret more complementary technologies as being more technologically advanced and productive than weakly complementary technologies (as in Fordism or lean manufacture versus more traditional sectors), then this extension provides a rationale (akin to Kremer (1993)) to differences in technological adoption that we observe empirically.[17] Finally, in Section 6.2 we show how the dynamic equilibria of our augmented model of Section 6.1 respond to changes in the underlying parameters of the model, to better comprehend why some regions have widely disseminated work ethic while others do not.

6.1 Firm Selection

In this section we augment our basic model to account for firm technological choice. We assume that firms can choose a technology $q \in \{C,\ S\}$, where C stands for the complementary technology presented in the base model, while S stands for a separable technology in which there are two projects, each one using a single worker and with success probability θa_j^{α}.[18] We again assume that the projects have no value if unsuccessful, but now their value in case of success is $\gamma\rho$. We interpret γ as the relative productivity of the complementary and separable sectors of the economy. If we portray those sectors as representing more modern and traditional sectors of activity, respectively, then γ corresponds to the relative productivity between them. We would

that, note that the derivative of the agent's utility in regards to *proposed* effort is:

$$u_1(a_i,\ a_j;\kappa_i) = \frac{c''(a_i;\ \kappa_i)y_1(a_i,\ a_j) - c'(a_i;\ \kappa_i)y_{11}(a_i,\ a_j)}{y_1(a_i,\ a_j)}y(a_i,\ a_j),$$

which is positive under the very standard assumptions that y is increasing and concave and c is increasing and convex.

[17] In this section our argument is similar to Kremer (1993), but, besides our focus on culture, we study effort choice and endogenize compensation choice by the firms, while their model is based on sorting on skill.

[18] Perhaps a more intuitive choice would be to use as probability of project success: $\theta\left(\sum_k a_k\right)^{\alpha}$. In this case, however, a_i and a_j wouldn't be separable, as we desire, but substitutes. The optimal effort proposal would then depend on the trade-off between this substitution effect and the convexity of the cost function. In the quadratic case, for example, since the marginal cost is linear, the expense of inducing effort a of two agents is the same as inducing effort $2a$ of one. In this case, clearly the firm will prefer to request positive effort only for the agent with smaller cost parameter κ_i.

expect then that events like the Industrial Revolution would change γ (namely, decrease it), and thus have an effect on the firm (and indirectly on the parental) choice. We also assume a lump sum cost of having complementary technologies ζ, which represents them being more modern and expensive production processes (both to acquire and to maintain).

In the separable technology, the utility the agent receives is independent of the effort (or type) of his co-worker, and is given by:

$$(\theta a^{\alpha})\, b_i - \kappa_i \frac{a^2}{2} \tag{7}$$

Our choice of separable technology is particularly appealing because it gives rise to exactly the same optimal expected compensation $\mu = \frac{\kappa_i a^2}{\alpha}$. In this case, however, since there are no complementarities, the effort proposed to an agent of type i depends only on his own type. It is given by:

$$a_i = \left[\frac{\rho\gamma\theta\alpha^2}{2\kappa_i}\right]^{A} \tag{8}$$

We denote the proportion of firms with complementary technology in the economy by β^C, and we assume that competition among firms causes low-achieving ones to be replaced by new entrants. Specifically, we consider the firms' fitness as proportional to their revenue relative to the mean revenue in the society; we model the selection process as a (discrete) replicator dynamic with a speed-scaling parameter s. Denoting by $Y(\beta_t^C,\ x_t^W;\ q)$ the expected revenue (net of complementary process cost ζ) of a firm with technology $q \in \{C, S\}$ at period t, the proportion of firms with complementary technology in period $t+1$ is then given by:[19]

$$\beta_{t+1}^C - \beta_t^C = s\beta_t^C \left[Y(\beta_t^C,\ x_t^W;\ C) - \sum_{q\in\{C,\ S\}} \beta_t^q\ Y(\beta_t^C,\ x_t^W;\ k) \right] \tag{9}$$

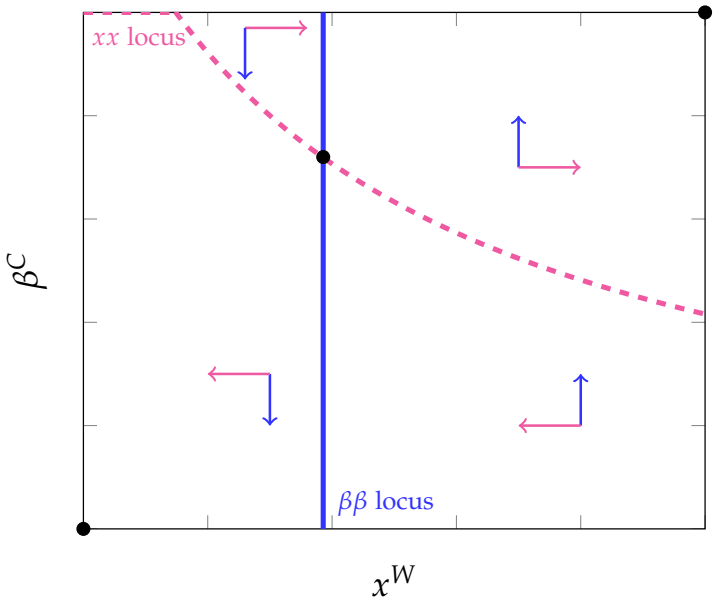
The existence of firms with separable technology also affects the parental socialization decision. Now parents must take into account that while with probability β^C their offspring will work in firms with complementary technology and receive payoff given by (5), with probability $1-\beta^C$ they will work in a firm without complementarities and receive $V^S(i)$, whatever their co-worker's type, where $V^S(i)$ is the utility of the agent, given by (7), evaluated at the optimal

[19] We use revenue instead of profits as selection process for tractability, since the profit is a highly nonlinear function of the agents' types, and clarity of interpretation. Nonetheless, under some conditions the results are still valid for dynamic processes based on selection on profit. The conditions under which our results will be true in that case are available on request.

contract. The agent's ex-ante labor market utility is thus given by:[20]

$$U^k(x^W,\ \beta^C) = \beta^C \left[x^W V^C(k,\ W) + (1 - x^W)V^C(k,\ N)\right] + (1 - \beta^C)V^S(k) \tag{10}$$

Figure 4: Dynamic system of x^W and β^C for an example parameter profile. ($\theta = 0.25$, $\rho = 10$, $\alpha = 0.7$, $\kappa_W = 0.4$, $\kappa_N = 0.8$, $\gamma = 0.55$, $\zeta = 0.8$, and $K_W = 0.45$)



Equations (9) and the cultural transmission process implied by (10) together generate a two-dimensional (discrete) dynamic system, of which we can study the continuous approximation by use of a phase space. This is done in Figure 4. The xx-locus is the set of points $(x^W,\ \beta^C)$ for which the proportion of agents with work ethic (x^W) is constant. Analogously, $\beta\beta$-locus is the set of points in the plane for which the proportion of firms with complementary technology β^C is constant. Notice that the loci include the borders, as both the replicator dynamics and the cultural transmission process do not incorporate mutations. The intersections between the loci represent stationary points of the dynamic system, of which $(0,\ 0)$ and $(1,\ 1)$ are stable, and the interior stationary point is a saddle-point.

Intuitively, when the proportion of agents with work ethic in the society is low, firms expect to employ few hard-working agents. Given that in modern complementary technologies shirkers will hinder the productivity of their peers, firms will try to avoid that by splitting them in separable production processes. On the other hand, when work ethic is highly disseminated, firms benefit from combining workers' effort in complementary tasks, as the probability of having

[20]Throughout this section we refer to the matching payoff defined in Section 4, $V(i,\ j)$, as $V^C(i,\ j)$, to emphasize that it is the payoff of an agent i (with co-worker j) working in a firm with complementary technology.

hard working employees is higher. As Henry Ford learned in Fordlândia, while a production system where each worker's effort is highly dependent of his peers might result in large gains of productivity when employees have work ethic, in the opposite environment it can be disastrous.

We formalize the discussion above in the Proposition 6.1 below. It states the condition under which the behavior of the extended model is well represented by Figure 4 above, that is, we have two stable equilibria, one with low dissemination of work ethics and separable technology, the other with high dissemination of work ethics and complementary technology.

Proposition 6.1. About the dynamic behaviour of the model with firm sorting:
(i) The $\beta\beta$-locus is a vertical line in the $x^W \times \beta^C$ plane, such that for lower levels of x^W the proportion of firms with complementary technology β^C is decreasing, and for higher levels of x^W it is increasing.
(iia) If $V^S(W) > V^S(N)$, then the only asymptotically stable dynamic equilibrium is with full dissemination of work ethic.
(iib) Otherwise the xx-locus is downward sloping and it crosses the $\beta\beta$-locus at at most one interior point. There are always two asymptotically stable equilibria, one with full dissemination of work ethics and in which all firms choose complementary technologies, and one where no individual has work ethic and all firms prefer the separable technology. If the xx-locus and the $\beta\beta$-locus intersect in an interior point, then there is also another dynamic equilibrium (at their intersection), which is a saddle-point.

The condition necessary for the behavior of this system to be similar to Figure 4 is that $V^S(W) < V^S(N)$. But notice that if it is indeed superior to have work ethic even when the technology is separable (i.e., $V^S(W) > V^S(N)$), then the dissemination of work ethic will always be high, with firms choosing to change to complementary technologies only amplifying this effect, and therefore the phase space would not be as Figure 4. Nevertheless, this would be contrary to the view (presented in Section 2) that work ethic was not already present in Medieval Europe, having instead emerged during the Protestant and Industrial Revolutions.

6.2 Comparative Dynamics

In this section we discuss how the dynamic equilibria are affected by changes in the underlying parameters. Since the cultural transmission process and the replicator dynamics of firm selection do not have mutations, in any stationary point parameter changes do not have any effect.[21] We can however study how changes in parameters affect the sizes of the basins of attraction for the

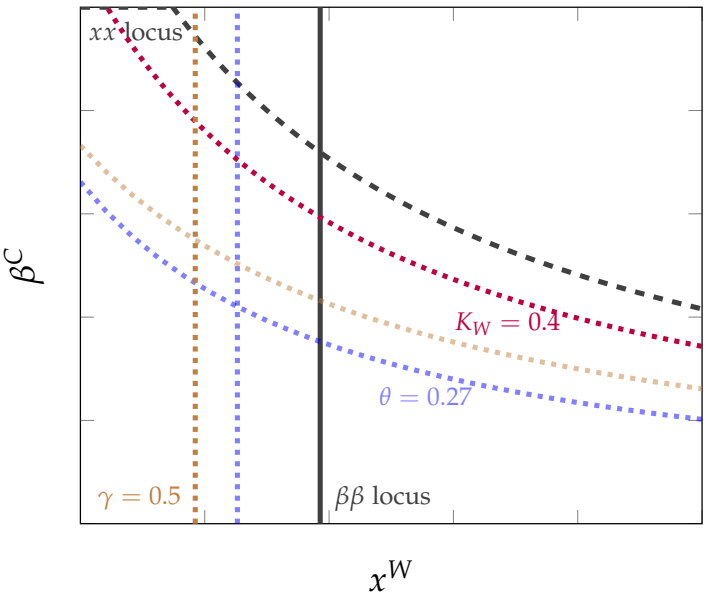
[21] We could augment the model with random shocks to the proportion of workers with work ethic. In that case, changes in parameters would alter the probability of the equilibrium changing. The intuition of the model would remain unchanged, however.

dynamic equilibria. These comparative dynamics exercises for the extended model of Section 6.1 are exemplified summarized in Proposition 6.2 below.

Proposition 6.2. Assume a phase space like Figure 4 (vide Proposition 6.1). Then an increase in human capital (θ) and value of project (ρ) and a decrease in cost of having work ethics (K_W) and in the relative productivity of the separable sector (γ) increase the basin of attraction of the equilibrium with high dissemination of work ethic and high proportion of firms with complementary technology (and *vice versa*).

Intuitively, when workers are more productive or the project they engage in is more valuable, firms will optimally choose to propose higher effort (vide (4)). But higher effort is less costly for individuals with work ethic than for those without, and so transmitting work ethic becomes more valuable for parents. Furthermore, since more productive workers exert more effort, it also becomes more valuable for the firm to choose modern complementary technologies, as the gains from complementarity are increased by higher level of effort. In this scenario, not only parents want to more strongly socialize their children into work ethic, but firms choose more complementary technologies and the equilibrium with disseminated work ethic and complementary technology becomes more likely (as in there are more initial states which converge to that point).

Figure 5: Dynamic system of x^W and β^C ($\theta = 0.25$, $\rho = 10$, $\alpha = 0.7$, $\kappa_W = 0.4$, $\kappa_N = 0.8$, $\gamma = 0.55$, $\zeta = 0.8$, and $K_W = 0.4$).



Moreover, a decrease in the fixed cost of having work ethic (e.g. moral condemnation for leisure) makes work ethic more desirable, while a reduction in the relative productivity of the separable sector (as technological advancements that rely on complementarities, like Fordism, or

that benefit disproportionally more modern sectors, like steam engines and computers) makes the firms more willing to adopt complementary technologies. Although the decrease in cost of having work ethic affects only the parental socialization choice directly, since the complementary technology is more attractive when the proportion of workers with ethic is higher, it also indirectly affects firm choice. Analogously, while changes in the relative productivity of the two technologies affect directly only the firms decision, workers with work ethic prefer to work in firms with complementary technologies (and *vice versa*), hence it also indirectly affects their choice.

Proposition 6.2 is illustrated in Figure 5, in which we consecutively apply changes to the parameters of Figure 3 (in grey). First, a decrease in the cost of having work ethic, K_W, due to religious or social support for toil, has the effect of lowering the xx-locus, increasing the basin of attraction of the hard-working equilibrium. Then, an increase in productivity (θ), due to increases in human capital, has the effect of both moving down the xx-locus and moving to the left the $\beta\beta$-locus, increasing both the initial proportion of individuals with work ethic and of firms with complementarity that lead to the equilibrium with full dissemination of both. Finally, lowering the relative productivity of the separable sector γ moves the $\beta\beta$-locus left, making the hard-working equilibrium more likely.[22] While these changes are purely illustrative, we believe they accurately represent changes that occurred in the 18th and 19th centuries in industrialized economies, leading to the dissemination of cultural norms of work values and the diffusion of production processes focused on complementarities, which we observe nowadays.

These comparative dynamics explain the impact of the Protestant Revolution, which disseminated an ascetic view of the importance of hard work for (religious) salvation, lowering the cost of effort and raising the cost of leisure. Consequently, it enlarged the basin of attraction of the equilibrium with hard-working agents, making work ethic more likely to disseminate. While it only affected the cultural transmission process directly, by allowing the dissemination of work ethic, the Protestant Revolution also permitted the adoption of more advanced technologies by firms, using the new hard-working population in more complementary production processes. Analogously, while the Industrial Revolution, by favoring modern production processes, induced the adoption of new technologies, by making effort more complementary, it also made work ethic more desirable, easing its dissemination. These two changes, therefore, had mutually reinforcing effects, creating a new equilibrium with high dissemination of work ethic and complementary technologies.

[22] A reduction of γ also has an effect of moving the xx-locus *up*, as it is a analogous to a decrease in ρ for the separable technology.

7 Conclusion

Cultural norms of work ethic have been studied in economics, but the focus so far has been mostly on the interaction between norms of work effort and political redistributional choices. In this paper we presented a model of dissemination of work ethic and incentive choice by firms, and we showed, in three separate empirical exercises, results that are consistent with this model.

While our model is not the only mechanism that is relevant in understanding cultural determinants of work behavior, it provides new insights and different policy implications from what has so far been established in this literature. For example, if one considers multiple equilibria in work ethic to be purely a political issue, then an exogenous change in tax rate would change the dissemination of work ethic in the society and move from one equilibrium to the other. We argue in this paper that firms' incentive steepness is related to work ethic, hence the previous assertion does not have to be true. There might be societies with the same level of redistribution and different levels of work ethic, because of differences in (endogenously chosen) firm incentives.

We further show that if we allow firms to choose their technology, and we interpret the complementary technology as a more modern one, then our model provides an explanation for why firms in third-world economies do not adopt more modern practices from advanced economies: given the mean level of effort in the economy, technologies in which the production of one worker is highly dependent of the production of his peers (e.g. when many workers manage a ring spinning textile machine) are less profitable than technologies in which workers production is independent (as in the putting-out system). The same argument can be made to enlighten why technological and productive advancement flourished in 18th Century Northern Europe, while it only reached Southern Europe in the end of 19th Century.

We explore more carefully this point by examining how the basins of attraction of the different equilibrium levels of work ethic change with changes in economic parameters. We find that both an increase in human capital and that a decrease in the cost of having work ethic, for example because of religious support to ascetic lifestyle, facilitate the expansion of both work ethic and the modern sector. We believe these comparative dynamics provide insights to the impact of higher literacy and the Protestant Ethic on the Industrial Revolution and overall economic development of 18th Century Northern Europe.

We hope future empirical research will further study the correlation between work ethic and firm compensation schemes. Particularly interesting would be studies that attempt to analyze the causal effect between those variables, searching for quasi-exogenous variations in both culture and incentives and checking their long-term effect.

References

Alesina, A. and G.-M. Angeletos (2005). Fairness and redistribution. *American Economic Review*, 960–980.

Baxter, R. (1996). *A Christian directory*. Soli Deo Gloria Publications.

Becker, S. O. and L. Woessmann (2009). Was weber wrong? a human capital theory of protestant economic history. *The Quarterly Journal of Economics 124*(2), 531–596.

Benabou, R. and J. Tirole (2006). Belief in a just world and redistributive politics. *The Quarterly Journal of Economics 121*(2), 699–746.

Bisin, A. and T. Verdier (2000). Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics 115*(3), 955–988.

Bisin, A. and T. Verdier (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory 97*(2), 298–319.

Bisin, A. and T. Verdier (2011). Chapter 9 - the economics of cultural transmission and socialization. Volume 1 of *Handbook of Social Economics*, pp. 339 – 416. North-Holland.

Boyd, R. and P. Richerson (1985). Culture and the evolutionary process.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics 117*(1), 339–376.

Brynjolfsson, E. and P. Milgrom (2012). Complementarity in organizations. *The Handbook of Organizational Economics 11*.

Cantoni, D. (2013). The economic effects of the protestant reformation: Testing the weber hypothesis in the german lands.

Cavalli-Sforza, L. L. (1981). *Cultural transmission and evolution: A quantitative approach*. Number 16. Princeton University Press.

Clark, G. (1987). Why isn't the whole world developed? lessons from the cotton mills. *The Journal of economic history 47*(01), 141–173.

Clark, G. (2008). *A farewell to alms: a brief economic history of the world*. Princeton University Press.

Doepke, M. and F. Zilibotti (2008). Occupational choice and the spirit of capitalism. *The Quarterly Journal of Economics 123*(2), 747–793.

Franklin, B. (2007). Advice to a young tradesman, written by an old one. *The Writings of Benjamin Franklin: Philadelphia, 1726-1757*.

Groshen, E. L. (1991). Sources of intra-industry wage dispersion: How much do employers matter? *The Quarterly Journal of Economics 106*(3), 869–884.

Hauk, E. and M. Saez-Marti (2002). On the cultural transmission of corruption. *Journal of Economic Theory 107*(2), 311–335.

Ichino, A. and G. Maggi (2000). Work environment and individual background: explaining regional shirking differentials in a large italian firm. *The Quarterly Journal of Economics 115*(3), 1057–1090.

Kremer, M. (1993). The o-ring theory of economic development. *The Quarterly Journal of Economics 108*(3), 551–575.

Milgrom, P. and J. Roberts (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica 58*(6), 1255–1277.

Stigler, G. J. and G. S. Becker (1977). De gustibus non est disputandum. *The American Economic Review*, 76–90.

Topkis, D. M. (1998). *Supermodularity and complementarity*. Princeton University Press.

Weber, M. (1998). *The Protestant ethic and the spirit of capitalism*. Roxbury Los Angeles.

Weibull, J. W. (1997). *Evolutionary game theory*. MIT press.

Appendix A

In Appendix A we prove all propositions stated in this paper. We ignore indifference situations throughout, as they have measure zero and would only encumber the analysis.

A.1: Proof of Proposition 4.1

Proof. Given a contract $\{w_{ij},\ b_{ij}\}$, the problem of an employee i matched with j is to maximize (1), which is a strictly concave problem (for $\alpha < 1$) with unique solution given implicitly by (3). Given (3), the firm maximizes (2), which is again a strictly concave problem for $\alpha < 1$ with unique solution given by (4).

Plugging this in the utility function of the agent, we get that:

$$V(i,\ j) = \kappa_i \left(\frac{1}{\alpha} - \frac{1}{2}\right) \left[\frac{\left(\theta\rho\alpha^2\right)^{1+\alpha A}}{2^{1+\alpha A}\kappa_i\kappa_j^{\alpha A}}\right]^{\frac{2A}{1-\alpha^2A^2}}.$$

All that remains to be shown is that the $V(i,\ j)$ above satisfies increasing differences in types. As it is well known, vide Topkis (1998), for C^2 functions increasing differences is equivalent to positive cross derivative. But indeed:

$$\frac{\partial V(i,\ j)}{\partial\kappa_i\partial\kappa_j} = V(i,\ j)\left(-\frac{2A\alpha A}{1-\alpha^2A^2}\right)\left(1-\frac{2A}{1-\alpha^2A^2}\right)\kappa_j^{-\frac{2A\alpha A}{1-\alpha^2A^2}-1}\kappa_i^{-\frac{2A}{1-\alpha^2A^2}-1} > 0$$

Remembering that $A \equiv \frac{1}{2-\alpha}$, all parameters are positive, and since $\alpha \in (0,\ 1)$, necessarily $A \in (\frac{1}{2},\ 1)$. All that remains to be shown then is that $\frac{2A}{1-\alpha^2A^2} > 1$. But clearly $2A > 1$ and $0 < 1-\alpha^2A^2 < 1$.

□

A.2: Proof of Proposition 4.2

Proof. For ease of reading we divide the proof in three steps. Throughout the proof, we assume that the initial state is interior, as otherwise the problem is trivial.

From equation (5), and substituting for the probabilities of success in socialization, we obtain that the parent of type i will choose $d^i \in [0,\ 1]$, such that:

$$d^i\left(x^i\right) = \text{argmax}_d\left\{U^j(x^j) + x^i\left[U^i(x^i) - U^j(x^j)\right] + d(1-x^i)\left[U^i(x^i) - U^j(x^j)\right] - C(d^i)\right\}$$

And thus he will choose d^i such that:

$$(1-x^i)\left[U^i(x^i) - U^j(x^j)\right] \leq d^i,$$

with equality if $d^i > 0$. Clearly, in equilibrium at each period only one kind of parent exerts positive socialization effort: the kind of parent for which the ex-ante labor market outcome is higher. That happens for W if and only if:

$$\Delta U(x^W) \equiv U^W(x^W) - U^N(x^W) = x^W\left[V(W,\ W) - V(N,\ W)\right] + (1-x^W)\left[V(W,\ N) - V(N,\ N)\right] > 0$$

According to Proposition 4.1, we know V is increasing differences in types, and thus either:
(a) $V(N,\ j) > V(W,\ j)$ for all j, and thus $\Delta U(x^W) < 0$ for all x^W;
(b) $V(W,\ j) > V(N,\ j)$ for all j, and thus $\Delta U(x^W) > 0$ for all x^W;
(c) $V(W,\ W) > V(W,\ N)$ and $V(N,\ N) > V(N,\ W)$.

But notice that rewriting $\Delta U(x^W)$, we obtain:

$$\Delta U(x^W) = [V(W, N) - V(N, N)] + x^W [(V(W, W) - V(N, W)) + (V(N, N) - V(W, N))]$$

In the case (c), $\Delta U(x^W)$ is always increasing in x^W, $\Delta U(0) < 0$ and $\Delta U(1) > 0$. Further:

$$\Delta U(x^W) = 0 \iff x = \frac{V(N, N) - V(W, N)}{(V(W, W) - V(N, W)) + (V(N, N) - V(W, N))}$$

Having established the conditions for which $d^i > 0$, it remains only to demonstrate the cultural transmission proceess converging properties. Indeed, since

$$x^W_{t+1} = x^W_t P^{WW}(x^W) + (1 - x^W_t) P^{NW}(x^W),$$

we obtain:

$$x^W_{t+1} = x^W_t (1 - x^W_t) \left[d^W_t - d^N_t \right],$$

which can be approximated by the continuous process $\dot{x}^W = x^W(1 - x^W) \left[d^W(x^W) - d^N(x^W) \right]$.[23] Thus, $\dot{x}^W > 0 \iff d^W(x^W) > d^N(x^W) \iff \Delta U(x^W) > 0$. □

A.3: Proof of Proposition 6.1

Proof. **Part 1:** First, for simplicity we normalize (with no loss of generality) the profit such that $Y^C_{WN} - Y^S_{WN} = 0$, where Y^q_{ij} is the probability of success of a firm with employees type i and j and technology q. Clearly since there are no mutations homogeneous populations are always stable. Naturally by (9), if $\beta^C_t \in (0,\ 1)$, then:

$$\beta^C_{t+1} > \beta^C_t \iff Y(\beta^C_t,\ x^W_t,\ C) > Y(\beta^C_t,\ x^W_t,\ S).$$

But by our assumptions this is independent from β^C_t. To show that the $\beta\beta$-locus is a vertical line, with β^C increasing on the right and decreasing on the left of it remains to be shown that $\Delta Y(\beta^C_t,\ x^W_t) \equiv Y(\beta^C_t,\ x^W_t,\ C) - Y(\beta^C_t,\ x^W_t,\ S)$ is always strictly increasing in x^W_t. Henceforth we drop the time subscripts for ease of notation.

Indeed, we have that:

$$\Delta Y(\beta^C,\ x^W) = (x^W)^2 \left[Y^C_{WW} - Y^S_{WW} \right] +$$

$$2x^W(1 - x^W) \left[Y^C_{WN} - Y^S_{WN} \right] + (1 - x^W)^2 \left[Y^C_{NN} - Y^S_{NN} \right]$$

[23] For details, see Weibull (1997).

Denote the derivative of $\Delta Y(x^W)$, shown above, by $D\Delta Y(x^W)$. Then it is (two times):

$$D\Delta Y(x^W) = \left[Y^C_{WN} - Y^S_{WN}\right] - \left[Y^C_{NN} - Y^S_{NN}\right] + x^W \left\{ \left[Y^C_{WW} - Y^S_{WW}\right] - 2\left[Y^C_{WN} - Y^S_{WN}\right] + \left[Y^C_{NN} - Y^S_{NN}\right] \right\}$$

Since the inside of the brackets is a constant (in x^W), the derivative is (generically) either strictly increasing or strictly decreasing in x^W, and it is sufficient to evaluate the equation above at the corners. A sufficient condition for the above derivative to be positive everywhere is that:

$$\left[Y^C_{WW} - Y^S_{WW}\right] > \left[Y^C_{WN} - Y^S_{WN}\right] > \left[Y^C_{NN} - Y^S_{NN}\right].$$

Under our normalization, that is equivalent to

$$\left[Y^C_{WW} - Y^S_{WW}\right] > 0 > \left[Y^C_{NN} - Y^S_{NN}\right].$$

There are therefore four possible cases:
(i) $Y^C_{WW} - Y^S_{WW} > 0$ and $Y^C_{NN} - Y^S_{NN} > 0$;
(ii) $Y^C_{WW} - Y^S_{WW} > 0$ and $Y^C_{NN} - Y^S_{NN} < 0$;
(iii) $Y^C_{WW} - Y^S_{WW} < 0$ and $Y^C_{NN} - Y^S_{NN} > 0$; and
(iv) $Y^C_{WW} - Y^S_{WW} < 0$ and $Y^C_{NN} - Y^S_{NN} < 0$.

We start by noting that (under our normalization) $\Delta Y(1) = Y^C_{WW} - Y^S_{WW}$ and $\Delta Y(0) = Y^C_{NN} - Y^S_{NN}$. Note also that $D\Delta Y(1) = Y^C_{WW} - Y^S_{WW}$ and $D\Delta Y(0) = -\left[Y^C_{NN} - Y^S_{NN}\right]$, and that $D\Delta Y(x^W)$ is a linear combinations of both. Using that, we now show that: a) case (iii) is impossible, and b) that in cases (i) and (iv) $\Delta Y(x^W)$ has no root.

a) This is direct result from $Y(\kappa_i,\ \kappa_j)$ being increasing differences in κ_i and κ_j. Indeed:

$$\frac{\partial Y^C_{ij}}{\partial i \partial j} = \left[-(1+\alpha A)\frac{\alpha A}{1-\alpha^2 A^2}\right]^2 C \kappa_i^{-(1+\alpha A)\frac{\alpha A}{1-\alpha^2 A^2}-1} \kappa_j^{-(1+\alpha A)\frac{\alpha A}{1-\alpha^2 A^2}-1} > 0,$$

where C is a constant (in $(\kappa_i,\ \kappa_j)$).

b) For (i), $D\Delta Y(1) \geq D\Delta Y(x^W)$ for all x^W (since they are a linear combination of $D\Delta Y(1) > 0$ and $D\Delta Y(0) < 0$). But, by the Fundamental Theorem of Calculus, if there was a y between 0 and 1 such that $\Delta Y(y) = 0$, then:

$$\Delta Y(1) = \Delta Y(1) - \Delta Y(y) = \int_y^1 D\Delta Y(x)\ \mathrm{d}x \leq (1-y)\Delta Y(1) < \Delta Y(1),$$

a contradiction. Analogously, for (iv), $D\Delta Y(x^W) \leq D\Delta Y(0)$, and thus if there is $y \in (0,1)$ such that $\Delta Y(y) = 0$, then we have *reductio ad absurdum*:

$$-\Delta Y(0) = \Delta Y(y) - \Delta Y(0) = \int_0^y D\Delta Y(x)\ \mathrm{d}x \leq \int_0^y D\Delta Y(0)\ \mathrm{d}x = y\left(-\Delta Y(0)\right) < -\Delta Y(0)$$

Therefore, either we have case (ii), or it is always superior to choose separable technology or always preferable to choose complementary technology, and the claim is proven.

Part 2: Now we examine the xx-locus. Since the ex-ante utility is given by (10), we have that:

$$\Delta U(x^W;\beta^C) = \left(V_W^S - V_N^S\right) + \beta^C \left\{x^W \left(V_{WW}^C - V_{NW}^C\right) + (1-x_W)\left(V_{WN}^C - V_{NN}^C\right) - \left(V_W^S - V_N^S\right)\right\}.$$

We now consider the hyperplane $\Delta U(x^W;\beta^C) = 0$ (which must be the case for any interior xx-locus point). First we note that for any given β^C, $\Delta U(x^W;\beta^C)$ is strictly increasing in x^W (because of Proposition 4.1). Thus we obtain that for any given β^C, if $\hat{x}$ is such that $\Delta U(\hat{x};\beta^C) = 0$, then for any $x^W > \hat{x}$, $\Delta U(x^W;\beta^C) > 0$ (and thus $x_{t+1}^W > x_t^W$, vide Appendix A.2) and vice-versa. Furthermore, since the derivative is everywhere positive, we can apply the Implicit Function Theorem on the surface $\Delta U(x^W;\beta^C) = 0$, and thus it must be true that:

$$\frac{\partial \beta^C}{\partial x_W} = -\frac{\partial_x \Delta U(x^W)}{\partial_\beta \Delta U(x^W)} = -\frac{\beta^C\left[\left(V_{WW}^C - V_{NW}^C\right) - \left(V_{WN}^C - V_{NN}^C\right)\right]}{x^W\left[\left(V_{WW}^C - V_{NW}^C\right) - \left(V_{WN}^C - V_{NN}^C\right)\right] + \left(V_{WN}^C - V_{NN}^C\right) - \left(V_W^S - V_N^S\right)}.$$

Since the numerator is always positive, for the surface to be downward sloping we need the denominator to be positive, that is for:

$$x^W\left[\left(V_{WW}^C - V_{NW}^C\right) - \left(V_{WN}^C - V_{NN}^C\right)\right] + \left(V_{WN}^C - V_{NN}^C\right) - \left(V_W^S - V_N^S\right) > 0,$$

But this is only relevant where $\Delta U = 0$. That happens when:

$$x^W\left(V_{WW}^C - V_{NW}^C\right) = \frac{(1-\beta^C)}{\beta^C}\left(V_N^S - V_W^S\right) + (1-x^W)\left(V_{NN}^C - V_{WN}^C\right)$$

Of course that is not valid for $x^W = 0$ or $\beta^C = 0$. Therefore the points x^W (which we denote as $\hat{x}$) for which $\Delta U = 0$ as a function of β^C are:

$$\hat{x} = \frac{\left(V_{NN}^C - V_{WN}^C\right) + \frac{(1-\beta^C)}{\beta^C}\left(V_N^S - V_W^S\right)}{\left(V_{WW}^C - V_{NW}^C\right) - \left(V_{WN}^C - V_{NN}^C\right)}$$

So the condition is that for all x satisfying this, that:

$$\frac{\partial \Delta U(x^W;\beta^C)}{\partial \beta^C} = \hat{x}\left[\left(V_{WW}^C - V_{NW}^C\right) - \left(V_{WN}^C - V_{NN}^C\right)\right] + \left[\left(V_{WN}^C - V_{NN}^C\right) - \left(V_W^S - V_N^S\right)\right] > 0$$

$$\iff \frac{(1-\beta^C)}{\beta^C}\left(V_N^S - V_W^S\right) + \left(V_N^S - V_W^S\right) > 0$$

Which always happens when $V_N^S > V_W^S$, and our proof is complete. □

A.5: Proof of Proposition 6.2

Proof. We show that $\frac{\partial[V(W,\,j)-V(N,\,j)]}{\partial\xi} > 0$, for $\xi \equiv \theta\rho$ (i.e. V is increasing differences in the agent's type and the parameters). The other two assertions are trivial, since they impact directly the relative payoffs of the different types. But indeed:

$$V(W,\,j) - V(N,\,j) = \xi^{(1+\alpha A)\frac{2A}{1-\alpha^2A^2}}\left[F(\kappa_W) - F(\kappa_N)\right],$$

where

$$F(\kappa_i) \equiv \kappa_i\left(\frac{1}{\alpha} - \frac{1}{2}\right)\left[\frac{\left(\alpha^2\right)^{1+\alpha A}}{2^{1+\alpha A}\kappa_i\kappa_j^{\alpha A}}\right]^{\frac{2A}{1-\alpha^2A^2}}$$

is an everywhere decreasing function of κ_i (since $\frac{2A}{1-\alpha^2A^2} > 0$) and thus the derivative is everywhere positive, as $\kappa_W < \kappa_N$. □

Appendix B (not for publication)

In Appendix B we provide descriptive statistics for the most important variables in all three empirical exercises. The tables are presented in the order of the main text, starting with tables for the International Social Survey data, followed by the European Values Study and European Social Survey data and the World Values Survey and Execucomp empirical exercise.

Table 4: Summary Statistics (ISSP 1989, 1997)

	Min	1st Quartile	Median	Mean	3rd Quartile	Max	Std. Deviation
Work important	0.2313	0.3915	0.5094	0.5286	0.6379	0.8668	0.1569
Education (yrs)	0	9	11	10.96	13	40	3.8461
Age	14	30	42	44.19	57	99	17.0257

Table 5: Summary Statistics for Binary variables (ISSP 1989, 1997)

	Home exhausted	%	Work Important	%	Female	%
Yes	25355	0.6000	24726	0.5282	25858	0.5299
No	16903	0.4000	22086	0.4718	22941	0.4701

Table 6: Summary Statistics (EVS 2008, ESS 2004, 2010)

	Min	1st Quartile	Median	Mean	3rd Quartile	Max	Std. Deviation
Work Important	0.0000	0.1958	0.4560	0.4706	0.7561	0.9661	0.2948
Age	14	32	47	47.76	62	102	18.7287
Education (yrs)	14	30	42	44.19	57	99	4.0480
Left right	0.0000	0.1958	0.4560	0.4706	0.7561	0.9661	0.5003
Bribe justifiable	0.0000	0.1843	0.2508	0.2898	0.3859	0.9444	0.1479
Concern for others	0.0000	0.1597	0.2166	0.2513	0.2887	0.7207	0.1421
Freedom	0.0000	0.4420	0.5311	0.5152	0.5945	1.0000	0.1036
log GDP pc	9.08	10.08	10.42	10.36	10.62	11.46	0.4263
Indústria (%)	15.74	25.37	28.89	29.22	31.95	45.22	6.0740

Table 7: Summary Statistics for Binary variables (EVS 2008, ESS 2004, 2010)

	Female	%	Wage depends on effort	%	East Europe	%	Self-employed	%
Yes	48440	0.5429	10037	0.2892	46088	0.5160	8826	0.1160
No	40789	0.4571	24672	0.7108	43227	0.4840	67270	0.8840

Table 8: Summary Statistics (WVS 1995, 1999, 2005, Execucomp 1994-2006)

	Min	1st Quartile	Median	Mean	3rd Quartile	Max	Std. Deviation
Work imp. (region)	0.4167	0.5476	0.5897	0.5829	0.6264	0.6735	0.0512
Bonus	0	0	0.2953	0.2847	0.4678	1	0.2428
Not salary	0	0.501	0.682	0.635	0.812	1	0.2323
Age	27	47	53	53	58	94	8.2072

Bonus is $\frac{\text{BONUS}}{\text{BONUS+SALARY}}$, the dependent variable in columns (1)-(3) and (6) in Table 3. Not salary is $1 - \frac{\text{SALARY}}{\text{TOTAL WAGE}}$, the dependent variable in columns (4) and (5).

Table 9: Summary Statistics for Binary variables (WVS 1995, 1999, 2005, Execucomp 1994-2006)

	Female	%	CEO	%	CFO	%	INTERLOCK	%
Yes	3214	0.0529	9559	0.1574	2913	0.0480	1038	0.0171
No	57500	0.9471	51155	0.8426	57801	0.9520	59676	0.9829