Teaching by example and induced beliefs in a model of cultural transmission*

Fabrizio Adriani † Jesse A. Matheson‡ Silvia Sonderegger§

July 28, 2017

Abstract

We augment standard models of cultural transmission with an explicit account of social learning, grounded in the information transmission literature. Youngsters observe the behavioral trait of a role model and form beliefs about the desirability of that trait. Adults have better information about each trait and have a paternalistic attitude toward their children. This makes them reluctant to adopt myopic behavior to avoid setting a negative example to their children. This signaling distortion increases in the influence parents have over their offspring. We extend the model to allow parental influence to depend on the population frequency of each trait and show that cultural complementarity does not imply convergence to a homogeneous population. We find empirical support for a positive relationship between parental influence and propensity to exert self-restraint by looking at alcohol and tobacco consumption.

JEL codes: D03, D82, D83, Z13. **Keywords: Parent-Child Signaling, Cultural Transmission, Social Learning, Conformity.**

*We thank an associate editor and two anonymous referees for suggestions. The paper has been partly inspired by discussions with A. Bisin, whom we thank. We have also benefited from comments by A. Barr, S. Bose, S. Currarini, V. Denicoló, J. Eguia, D. Myatt, L. Samuelson, C. Wallace, and P. Zanchettin.

†University of Leicester. Department of Economics, University road, Leicester, UK, LE17RH. e-mail: fa148@le.ac.uk, phone: +44(0)1162523371 (corresponding author).

‡University of Leicester. Department of Economics, University road, Leicester, UK, LE17RH. e-mail: jm464@le.ac.uk, phone: +44(0)116 252 2494.

§University of Nottingham and CeDEx, e-mail: silvia.sonderegger@nottingham.ac.uk, phone: +44(0)115 84 67347.

1 Introduction

The cultural transmission literature has roots in the pioneering work of socio-biologists such as Cavalli Sforza and Feldmann (1981) and Boyd and Richerson (1985). Starting from Bisin and Verdier (1998, 2001a), a literature on this topic has also recently emerged in economics (see Bisin and Verdier, 2011, for a survey of contributions). These works have augmented the early socio-biology models with key elements of economic theory like strategic behavior and optimization, thus firmly setting cultural transmission within the boundaries of standard economics. The core features of most existing models, however, are still to a large extent borrowed from models of genetic transmission. Cultural transmission is seen as a sort of "black box". Young individuals typically acquire their cultural traits from role models (who may be their parent or another adult), but the way in which this socialization process works is not explicitly spelled out. We argue that this is unnecessary. Economists have long worked with models of information transmission and (Bayesian) social learning that have proved useful in understanding important socio-cultural phenomena (see e.g. Banerjee, 1992, and Bikchandani et al., 1992).

The idea that social learning may matter for cultural transmission is of course not new (see e.g. Bandura and Walters, 1963). Many existing models of cultural transmission acknowledge the importance of social learning and often justify their assumptions by invoking it.[1] However, there is almost an implicit belief that augmenting existing models with a worked-out account of social learning would add little, if anything at all. This implicit understanding has not been properly scrutinized, though.

This paper aims at filling this gap. We build a simple model, where agents are heterogeneous in their tastes and can adopt one of two alternative behavioral traits (e.g. self-gratification or abstinence). In our model, there is thus a conceptual distinction between tastes (that agents take as given) and behavioral traits (that agents can intentionally adopt). We then use this setup to pit the predictions of standard mechanisms of cultural transmission against those of a more sophisticated mechanism which also includes social learning. In

[1] For instance, Bisin and Topa (2010) mention *indirect socialization*, whereby "...children ... imitate or learn from their parents' or siblings' behavior", as an important aspect of socialization. Moreover, social learning is, together with genetic inheritance, one of the pillars of the dual inheritance theory (see Cavalli-Sforza and Feldman, 1981, Boyd and Richerson, 1985).

standard models, the process of cultural transmission takes the form of children passively inheriting their preferences from a role model of the older generation. The social learning model augments this mechanism by allowing also for information transmission: Children observe the behavior of their role model and update their beliefs about the desirability of the observed behavioral trait.

We find that explicitly accounting for the role of social learning generates novel effects. The key observation is that the adults' behavior conveys information to children, thus generating an information externality. Moreover, informed parents are aware that their own behavior sets relevant "examples" to their children, and may thus modify it strategically. This feature is absent from standard models, since the children's behavior is usually entirely determined by their inherited preference trait.

In principle, information transmission need not by itself affect outcomes. However, it becomes relevant when parents have a paternalistic attitude towards their offspring, so that their objectives are imperfectly aligned with those of their children. In this case, information transmission leaves a signature in the form of a "signaling distortion" in parental behavior. To see how this works, consider for instance a parent concerned that her child may take up smoking. Acknowledging that this may be setting a "bad example" to the child (e.g., by suggesting that smoking is not harmful, or that it is the socially acceptable thing to do), she may refrain from smoking herself. What's more, her incentive to quit smoking will be larger the larger her influence (in the sense described above) over the child. Hence, the social learning model predicts a positive relationship between parental influence and *the parents'* propensity to exert self-restraint. We argue that this "disciplining effect" can be used to empirically validate the role of information transmission in socialization.[2]

Recent experimental and field evidence is consistent with the hypothesis that parents do modify their behavior when observed by their children. In a coin-toss experiment, Houser

[2]In standard models, the cultural transmission mechanism exclusively works at a deeper level, in the domain of *preferences*. As a result, there is no incentive to distort one's behavior in order to strategically manipulate information transmission. Of course, one could always postulate that children's preferences are shaped by the *behavior* of their role models (as opposed to their preferences), thus replicating the main feature of the social learning model. However, we argue that such a theory would be more parsimoniously captured by a model of social learning where agents behave *as if* they were strategically sending (and receiving) information about the actual consequences of a particular behavior.

et al. (2016) find that parents are more reluctant to cheat when their child is in the room. Similarly, Ben-Ner et al. (2015) report that parents increase their contributions in the dictator game when these will be shown to their children.

In a dynamic extension, we also look at the implications of the discipling effect for the long run distribution of traits in the population. In their seminal contribution, Bisin and Verdier (2001a) analyze the forces that may generate the persistence of the high degree of cultural heterogeneity observed in modern societies. They identify *cultural substitutability* as a key factor.[3] Under cultural substitutability, as a cultural trait becomes rarer, the influence of parents displaying the rare trait must increase relative to that of parents displaying the alternative trait. We show that social learning generates richer dynamics and that some form of heterogeneity may persist even in the presence of cultural complementarity. The intuition is as follows. Suppose that, as a behavioral trait involving abstinence spreads among the population, the influence of adults displaying the alternative trait (e.g. smoking) *is reduced*; i.e. smoking adults become relatively less likely to transmit their trait to the next generation. This type of cultural complementarity might for instance arise if the signals sent by parents become less salient when they do not conform to prevailing behavior in society. Under standard cultural evolution, if the initial share of adults choosing abstinence is large enough, abstinence would then spread all the way and the alternative trait would vanish in the long run. With social learning, in contrast, the behavioral trait adopted by parents is endogenous with respect to parental influence. As the influence of parental smoking fades, adults who like smoking become less concerned about the example set to their children. Some of those who would choose abstinence for signaling reasons may thus conclude that it is safe to switch to smoking. This effect clearly works as a counterweight to the spread of abstinence, thus allowing for the persistence of some degree of heterogeneity.

A secondary benefit of explicitly distinguishing between intrinsic tastes and behavioral traits is that the model is sufficiently rich to also accommodate the social rewards from conforming to the behavior of peers or to the norms of society at large.[4] We thus study

[3]See also Panebianco and Verdier (2015) for how cultural substitutability preserves heterogeneity even when the transmission network is biased toward a specific trait.

[4]The role of conformity is also acknowledged by socio-biologists, who typically allow cultural transmission to have a *conformist bias* See also Saéz-Martí (2011) and Saéz-Martí and Zenou (2012). Bernheim (1994) provides an alternative way to model pressure to conform.

how the desire to conform interacts with parents' signaling distortion to shape observed behavior and culture. For instance, a childless individual may refrain from smoking not in order to set an example to his (non-existing) child, but because his acquaintances, who may have children, are non smokers. In practice, most people form their habits in part by interacting with others. Working, eating, drinking, smoking, religious rituals are partly social activities. From an empirical viewpoint, this complicates the task of detecting the effect of social learning by using differences in parental influence across individuals, who may in fact share the same norms of behavior. Whether individual parental influence is a good predictor of adults' propensity to adopt a particular habit or the desire to conform tends to prevail is essentially an empirical question.

In Section 6 we provide some evidence for the key empirical implications of model. We use data collected as part of the Iowa Youth and Families Project. These data cover 451 households and include a parental influence index based on the reported influence that both fathers and mothers believe they have on their children. This information makes the data uniquely suited to assess the relationship between parental influence and parents' propensity for self-restraint. We compare smoking and drinking habits across parents with low and high perceived influence, controlling for a number of parent, child and household characteristics. We find strong evidence that high influence fathers refrain more from smoking than low influence ones. The results for drinking are more noisy but go in the same direction. Interestingly, while both high influence mothers and high influence fathers tend to drink less than their low influence counterparts, the results are statistically significant only for mothers. To better account for heterogeneity across households, we also use propensity score matching to estimate the effect of parental influence by matching households that are homogeneous in terms of observable characteristics. Relative to the results obtained with OLS, we find significant differences in smoking and drinking behavior both for fathers and mothers in the direction predicted by the theory. Finally, we check the reliability of our approach by looking at a type of behavioral outcome for which the theory *does not* predict that high influence parents should show more restrain than low influence ones. To this purpose, we use information on church attendance. Consistent with intuition, the data show that high influence parents *are not* less likely to attend church (in fact church attendance is increasing with perceived influence, although not in a statistically significant way when enough controls are added).

In order to understand the implications of our results for children's outcome – and to further validate the model in terms of cultural transmission – we match parental data with the child's propensity to experiment with smoking. In line with the theory and with existing evidence, we find that children of smoking parents are more likely to try smoking than other children. Of course, we cannot say how much of this correlation is due to example setting by parents and how much is driven by unobserved factors. However, it is interesting to note that children of high influence parents are less likely to experiment with smoking. Furthermore, this effect is entirely driven by children who have been "vertically socialized" in the sense that they report wanting to be like their parent.

Gaining a better understanding of the cultural transmission mechanism is not purely an academic exercise. Policy proposals like imposing age restrictions on all films portraying people engaged in the consumption of alcohol or tobacco ostensibly aim at manipulating cultural transmission.[5] Since these policies are unlikely to affect the physical pleasure of (i.e. the deep preferences for) consuming alcohol or tobacco, their potential effects can only be understood by explicitly analyzing information transmission. More generally, our theory points to reducing the separation of adults and children as a potential policy lever to encourage more farsighted behavior. The current lifestyle of most people in advanced economies dictates that children spend most of their time at school and parents spend most of their time on the workplace. This segregation may take extreme forms, as for instance in the case of Chinese rural immigrant workers – and their "left behind" children who are left in the care of grandparents and other relatives whilst the parents move to urban areas in seek of employment opportunities.[6] Most of the debate on the "left behind" generation has focused so far on the effects on children's outcomes and well being. Our analysis suggests that the effects on parents of a reduction of parental influence should also be given careful consideration.

The paper is organized as follows. After briefly reviewing the theoretical literature, we present in Section 2 a two stage version of the model which illustrates the static effects

[5] See Longacre et al. (2009).

[6] The China Daily recently reported that the number of "left behind" children under 17 in rural areas exceeds 60m.

at work. Section 3 provides a full characterization of the equilibrium. Section 4 presents an OLG version of the model and analyzes the long run distribution of behavioral traits. Section 5 returns to the two stage version of the model to devise empirical predictions and analyze the role of conformity. Section 6 presents some evidence suggesting a relationship between parental influence and self-restraint and discusses the relevant empirical literature. Concluding remarks are presented in Section 7.

1.1 Related theories

There is a small but growing economic literature on parenting, which partially overlaps with the literature on cultural transmission (see e.g. Weinberg, 2001, Linbeck and Nyberg, 2006, Lizzeri and Siniscalchi, 2008, Corneo and Jeanne, 2009, Doepke and Zilibotti, 2011 and 2014). This literature has however focused on trade-offs that do not directly emerge from the problem of setting an example to children and is thus complementary to the present work. In particular, Lizzeri and Siniscalchi (2008) consider the trade off between sheltering the child from mistakes and encouraging learning from experience. In a similar vein, Doepke and Zilibotti focus on the transmission of risk preferences (Doepke and Zilibotti 2011) and parenting styles (Doepke and Zilibotti 2014).[7] A common theme in these contributions is the trade off between direct parental intervention (which we may call *parental control*) and child's autonomy. It is uncontroversial that in reality parents also try to prevent their children from engaging in myopic or hazardous behavior through a mix of monitoring and discipline. We abstract from these tools in order to focus on teaching by example, but these could be easily incorporated into our theory. For instance, some parents induce the desired behavior through teaching (including teaching by example), while respecting the child's autonomy (*authoritative parenting*). Other parents rely more on direct control (*authoritarian parenting*). So long as influence (in the sense above) and control are substitutes, allowing for direct parental control would magnify the effect we highlight. As more influence tends to be associated with less control, the example set to children by high influence parents becomes more important.

[7] Weinberg (2001) focuses on the choice between the use of an authoritarian parenting style and monetary incentives. Lindbeck and Nyberg (2001) consider an incentive/insurance trade off. Corneo and Jeanne (2009) look at the trade off between giving children an open mind and making them proud of their cultural traits.

Some of the key ingredients of our analysis are also present in Bénabou and Tirole (2006). As in our case, parents strategically manipulate the information available to children in order to partially redress a time inconsistency problem. In their model, this is however obtained by suppressing information, while we focus on example setting. As a result, their model does not predict any relationship between parental influence and observable parental behavior. A similar argument applies to Cervellati and Vanin (2013), where parents induce children to exert self-restraint by endowing them with moral values. Pichler (2010) considers a preference transmission model where adults distort their behavior to affect their children's preferences. In our model, in contrast, adults' behavior aims at manipulating the *beliefs* of children. Finally, in a companion paper (Adriani and Sonderegger, 2016) we consider a framework similar to the one adopted here. That work also looks at the effects of parent-child signaling, but the signaling distortion is of a different nature. In that case, it is driven by the presence of strategic uncertainty and the parents' desire to shed their children from the costs of miscoordination. Here, we focus instead on the simpler – and perhaps more empirically relevant – case of a distortion resulting from a conflict of interest between adults and their children.

2 A two stage model

We start off by analyzing a two stage version of the model to better highlight the static effects at work. The environment is a two stage binary action global game (Carlsson and van Damme, 1993) where adults adopt a behavioral trait/behavior in the first stage and youngsters choose what trait to adopt in the second. Agents are heterogeneous in their preferences over a given behavior. While each adult knows her own tastes, the distribution of tastes in the population is unknown. Youngsters are uncertain about the desirability of each behavioral trait and infer it from the observed behavior of an adult role model. An OLG version of the model is presented in Section 4.

2.1 Environment

Behavioral traits and payoffs We consider a setup with two generations, $\tau = Y, A$, where Y stands for "youngsters" and A stands for "adults". There is a continuum of parent-child pairs indexed by $i \in [0, 1]$. Each agent, whether adult or youngster, can choose between two behavioral traits $a \in \{0, 1\}$. Action $a = 1$ (self-gratification) can be thought of as engaging in potentially myopic or hazardous behavior (e.g. smoking, drinking, crossing outside cross walks). Action $a = 0$ (abstinence) can be interpreted as abstaining from such behavior. We set the payoff from abstinence equal to zero. The payoff from $a = 1$ of individual $i \in [0, 1]$ of generation $\tau = Y, A$ is separable as follows

$$u_{\tau i} + v(x_\tau, x_s) \tag{1}$$

with $s = Y, A$, $s \neq \tau$. The term $u_{\tau i}$ represents individual τi's intrinsic taste for self-gratification, net of its costs. We assume that the taste term can be decomposed into $u_{\tau i} = \overline{\boldsymbol{u}} + \theta_{\tau i}$, where $\overline{\boldsymbol{u}} \in \mathbb{R}$ is a constant term common to all, and $\theta_{\tau i}$ is an individual specific taste shock. Individual taste shocks for the adult cohort are assumed to be uniformly drawn in the interval $[\theta - \epsilon, \theta + \epsilon]$. The distribution of tastes among the youngsters is determined by the preference transmission mechanism (see below).

The function $v : [0, 1]^2 \to \mathbb{R}$ measures the pressure to conform coming from society. The pair $\{x_\tau, x_s\} \in [0, 1]^2$ represent the share of individuals choosing $a = 1$ in generations τ and s, respectively. We assume that v is bounded, differentiable and, in order to induce a coordination motive, non-decreasing in both its arguments. In words, the utility derived from self-gratification is higher when most people do the same. Without any loss of generality, we set $\int_0^1 \int_0^1 v(x_\tau, x_s) dx_\tau dx_s = 0$.[8]

It is worth stressing that our model nests various settings, including: 1) A setup where the social rewards from self-gratification only depend on the total mass of individuals choosing it: $v(x_\tau, x_s) = \tilde{v}(x_\tau + x_s)$, 2) A setup where individuals are only concerned about conforming to members of their generation, i.e. $v(x_\tau, x_s) = \tilde{v}(x_\tau)$.[9]

Paternalism Adults are altruistic towards their own children and do not care about the welfare of other agents. Adults' total utility is thus given by the sum of their own utility and their perception of their child's utility. Similar to Bénabou and Tirole (2006) and Doepke and Zilibotti (2014), a crucial assumption in our setup is that, when evaluating the welfare of their children, parents are more farsighted than their children. In particular, we assume

[8] This can always be obtained by suitably reparametrizing the parameter $\overline{\boldsymbol{u}}$.

[9] It is also trivial to extend the model to accommodate asymmetric situations where, say, adults do not care much about coordinating with youngsters while youngsters want to coordinate with the adults.

that self-gratification entails long term costs $c > 0$, to be borne at an unspecified future date. When choosing their behavioral trait, agents discount the long term consequences for *their own* well being with a factor which, without loss of generality, we set equal to zero.[10] In contrast, when assessing *their children's* welfare, adults use a larger discount factor $\delta \in [0, 1]$. Hence, δ measures the degree of *paternalism*. We should note at this point that paternalism induces a potential conflict of interests between parents and children. There may exist situations where both a parent and her child prefer one action (e.g. smoking) but the parent wants her child to choose the other (e.g. not smoking).[11]

Role models and parental influence In the first stage, all adults privately observe their own taste term and simultaneously choose their behavioral trait. Then, each youngster is randomly assigned a role model. With probability $\lambda \in [0, 1]$, this is the youngster's parent. With probability $1 - \lambda$, it is another individual randomly drawn from the adult population. In the second stage, youngsters observe the behavior of their role model and select their behavioral trait.

The parameter λ, which is central to our analysis, thus measures the degree of *parental influence*. Keeping everything else equal, we expect λ to be high when children have high exposure to their parents, for instance because society is characterized by strong family ties.[12] By contrast, environments characterized by pervasive external influences (generated for instance by the media) that aggressively promote their lifestyle models correspond to a lower λ. It is worth noting that the literature tends to restrict the notion of *vertical transmission* to socialization occurring inside the nuclear family, while socialization by teachers, carers, other relatives or family friends is bunched together with socialization by, say, celebrities or fictional characters under the same *oblique transmission* label. In our setting, the key operational distinction between vertical and oblique transmission is the presence or absence

[10] If all agents had a positive discount factor, $\beta > 0$, one could always obtain the present model by replacing the term $\overline{\boldsymbol{u}}$ with $\overline{\boldsymbol{u}}^* \equiv \overline{\boldsymbol{u}} - \beta c$.

[11] An alternative way to interpret our model is that parents do not fully internalize their children's cost of exerting self-control. This is important since empirical evidence suggests that time discounting may not be the main driver of smoking behavior, while self control may be (see Khwaya et al. 2007).

[12] Alesina and Giuliano (2014) find that, although societies with strong family ties tend to have lower levels of economic activity and more traditional attitudes toward women, they also tend to report higher levels of health and well being. Our theory could potentially provide a non-obvious channel through which family ties may affect health/well being.

of concerns toward the youngster's welfare. For the purposes of our model, it is thus perhaps more convenient to think of socialization by all adults potentially concerned about a youngster's well being as more similar in nature to vertical transmission than to oblique transmission.

Various models of cultural transmission postulate (or endogenously generate) heterogeneity in λ (see e.g. Bisin and Verdier, 2001a). For the time being, we will assume that parental influence is exogenous and the same for all adults. These assumptions will be relaxed in Sections 4 and 5.1.

Preference transmission We follow existing literature and assume that each youngster Yi inherits his taste term from his role model Aj, so that $u_{Yi} = u_{Aj}$. As already mentioned, with probability λ, youngster i's role model is his parent (so that $Aj = Ai$). With the residual probability $1-\lambda$, Aj is a random adult different from Ai. This mechanism, which blends vertical and oblique transmission, is the backbone of most models of cultural transmission (e.g. Cavalli-Sforza and Feldman, 1981, Boyd and Richerson, 1985, Bisin and Verdier, 2001a). We retain this assumption to ease the comparisons with existing literature, although other mechanisms would generate identical results in our setting. In particular, genetic transmission (tastes inherited from the biological parent), conditional independence (youngsters' tastes independently drawn conditional on θ) and standard vertical/oblique transmission are all equivalent in the two stage version of the model – although the dynamics in Section 4 would need to be modified according to the chosen mechanism.[13]

Note that the nature of the transmission mechanism ensures that the distribution of θ_{Yi} is identical to the distribution of adults' taste terms θ_{Ai}, so that they are both uniform in $[\theta - \epsilon, \theta + \epsilon]$.

[13] Formally, our results apply to all settings where $\mathbb{E}[u_{Yi}|\theta_{Ai} = z] = \overline{\boldsymbol{u}} + z$. In words, a parent's expectation of her child's taste is equal to her own taste. While genetic transmission obviously satisfies this requirement, both the standard preference transmission framework and conditional independence are also consistent with it when adults do not know the mean θ of the population. Ignoring boundary problems, each adult Ai rationally expects the distribution of tastes of others to be centered around his own taste, so that $\mathbb{E}[\theta_{\tau j}|\theta_{Ai} = z] = \mathbb{E}[\theta|\theta_{Ai} = z] = z$. Hence, $\mathbb{E}[u_{Yi}|\theta_{Ai} = z] = \overline{\boldsymbol{u}} + \lambda z + (1-\lambda)\mathbb{E}[\theta_{Aj}|\theta_{Ai} = z] = \overline{\boldsymbol{u}} + z$. An equivalent interpretation of this condition, which we exploit literally in the dynamic version, is that adults evaluate youngsters' utility through the lenses of their own tastes as in Bisin and Verdier (2001a) and subsequent literature.

Information As mentioned above, adults and youngsters have different information about the desirability of self-gratification. It seems plausible that youngsters may be unable to accurately foresee the pleasure of casual sexual intercourse, or the addictiveness of smoking and drinking habits. We model this by assuming that, while all adults privately observe their own taste, u_{Ai}, youngsters do not observe u_{Yi} before choosing their action, but observe instead the behavioral trait of their role model. We also assume that the distribution of tastes is unknown. Agents do not observe the average taste shock θ, which is uniformly drawn in the interval $[-D, D]$.[14]

Since taste terms are correlated, the role model's behavior conveys a signal about the desirability of either behavioral trait. It is worth noting that all our results can be obtained in an equivalent version of the model where there is uncertainty about the (possibly heterogeneous) long term consequences of self-gratification, rather than tastes. In the text, we will maintain that the uncertainty is about tastes rather than consequences as tastes provide a more intuitive source of heterogeneity.

3 Analysis

Let Δp represent the difference between the probability that a youngster will engage in self-gratification after his role model has chosen self-gratification and the same probability when his role model chooses abstinence. If adult j is youngster i's role model, then

$$\Delta p \equiv \Pr(a_{Yi} = 1 | a_{Aj} = 1) - \Pr(a_{Yi} = 1 | a_{Aj} = 0). \tag{2}$$

The expected net payoff of a generic adult Ai from $a = 1$ can be expressed as

$$\mathbb{E}[v(x_A, x_Y)|\theta_{Ai}] + \overline{\boldsymbol{u}} + \theta_{Ai} + \lambda \Delta p \left[\mathbb{E}[v(x_Y, x_A)|\theta_{Ai}] + \overline{\boldsymbol{u}} + \theta_{Ai} - \delta c \right], \tag{3}$$

where we used the fact that, conditional on adult Ai being youngster Yi's role model, $u_{Yi} = u_{Ai}$.[15] The first three terms in (3) represent the adult's direct expected utility, while

[14] The assumption that the mean shock in the population is unobservable allows us to exploit some results in the literature on global games (see e.g. Morris and Shin, 2003).

[15] Note that the same result obtains whenever, conditional on vertical transmission, a parent *expects* his child to share his own taste, i.e. $\mathbb{E}[u_{Yi}|u_{Ai}] = u_{Ai}$. Since adults do not know the mean taste and use their own taste to infer it, this would for instance be the case if youngsters' and adults' tastes were independently drawn from identical distributions with unknown mean (conditional independence).

the term in brackets captures the consequences of parental behavior on the child's utility. The latter is discounted by the factor $\lambda\Delta p$, which is the probability with which the parent's choice will affect her child's choice.

It is instructive to compare the model with social learning with the case where children are informed about their tastes and there is no information transmission. The crucial difference between the two models lies in the different role played by Δp. The behavior of a youngster who is aware of his tastes is not affected by the behavior of his role model. As a result, when choosing between the two actions, a parent maximizes her utility taking Δp equal to zero. By contrast, in the social learning case, the role model's action conveys information about the youngster's tastes and thus affects the youngster's behavior. An information externality is thus present in this case. The value of Δp depends on the youngster's equilibrium strategy and will be generally different from zero.

Bringing forward some of the results, it is possible to show that, for D sufficiently large, youngsters always choose the same action as their role model in equilibrium independently of whether they are informed or not. However, with informed youngsters, the correlation is spurious. A youngster and his role model choose the same action because they happen to have perfectly correlated tastes. By contrast, in the social learning model, there is a causal relationship between the behavior of the role model and that of a youngster.

Before moving on to characterize the equilibrium of the game, it is worth discussing briefly the issue of multiplicity. Our setup allows both for payoff complementarities and for information externalities. Both complicate the task of pinning down a unique equilibrium. A standard feature of binary action global games is that the indeterminacy disappears when there is sufficient fundamental uncertainty about θ (see e.g. Morris and Shin, 2003). As shown below, this also eliminates any indeterminacy due to signaling in our context. In what follows, we will thus characterize the equilibrium for D sufficiently large. It is however worth stressing that a large D is mostly needed to establish uniqueness. The equilibrium we characterize exists so long as the fundamental uncertainty (D) is not too small relative to the degree of heterogeneity in preferences (captured by ϵ). In economic terms, assuming a large D relative to ϵ means that youngsters' prior information is relatively poor relative to the information available to adults – so that youngsters learn a lot from observing their role model's behavior. A potential drawback of this assumption is that the incentive to imitate one's adult role model trumps other considerations, like for instance conforming to peers. By relaxing this assumption, our model could be used to analyze how, as the information content of adults' behavior becomes weaker, youngsters may choose to depart from the norms of the previous generation.[16] However, since our focus here is on the behavior of adults and signaling concerns, we abstract from these issues.

3.1 Equilibrium characterization

Youngsters are aware of the preference transmission mechanism and thus understand that their tastes are correlated with those of the adults, but do not know ex-ante their own tastes. As a result, the role model's behavior conveys valuable information to the youngster.

Lemma 1. *For D sufficiently large, the youngsters' equilibrium strategy is to imitate their role model, i.e. $\Delta p = 1$.*

Proof. See supplementary material.

This implies that, in any equilibrium, $x_A = x_Y \equiv x^*$. Consider an equilibrium where adults choose $a = 1$ if and only if their private signal θ_{Ai} is above a cutoff (the same for all adults). In the supplementary appendix, we show that in this case $\mathbb{E}[v(x_\tau, x_s)|\theta_{\tau i}]$ is non-decreasing in $\theta_{\tau i}$ and that it is zero when evaluated at $\theta_{\tau i}$ equal to the cutoff. Imposing $\Delta p = 1$ in (3), one can then determine the cutoff θ^*_{SL} as the realization of $\theta_{\tau i}$ that makes an adult indifferent between the two actions,

$$\theta^*_{SL} = \frac{\lambda}{1+\lambda}\delta c - \overline{\boldsymbol{u}}. \tag{4}$$

Adults thus choose $a = 1$ whenever their taste shock θ_{Ai} is above θ^*_{SL} and youngsters fully imitate their role model. The observed share of individuals choosing self-gratification, x^*, depends on the value of θ^*_{SL}, which is determined by the parameters of the model, and on the average shock in the population, θ,

$$x^* = \begin{cases} 1 & \text{if } \theta > \theta^*_{SL} + \epsilon \\ \frac{1}{2} + \frac{\theta - \theta^*_{SL}}{2\epsilon} & \text{if } \theta^*_{SL} - \epsilon \leq \theta \leq \theta^*_{SL} + \epsilon \\ 0 & \text{if } \theta < \theta^*_{SL} - \epsilon \end{cases} \tag{5}$$

[16] A prominent example are attitudes toward same sex marriage in the US, which show a clear generational divide (see e.g. Sherkat et al., 2011.).

Clearly, the higher θ^*_{SL}, the lower the propensity to choose self-gratification. We thus define the cutoff as the *propensity to abstain*.

The next result, whose proof heavily borrows from Morris and Shin (1998), establishes that this equilibrium is the unique equilibrium for D sufficiently large.

Proposition 1. *For D sufficiently large, the social learning model has a unique equilibrium where each adult Ai chooses self-gratification $a = 1$ if and only if her taste term θ_{Ai} is above a cutoff θ^*_{SL} given by (4). Each youngster chooses self-gratification if and only if he observed his role model choosing the same. In equilibrium the share of adults x_A and the share of youngsters x_Y choosing self-gratification are both equal to x^* in (5).*

Proof. See supplementary material.

The comparative statics are summarized in the next corollary,

Corollary 1. *The propensity to abstain, θ^*_{SL}, and (for $x^* \in (0,1)$) the share of agents choosing abstinence, $1 - x^*$, are both decreasing in the (prior) average desirability of self-gratification $\overline{\boldsymbol{u}}$, and increasing in: the long term consequences of self-gratification, c, adults' paternalism, δ, and parental influence λ.*

As one might expect, self-gratification becomes more widespread as agents enjoy it more (higher $\overline{\boldsymbol{u}}$), and less widespread as its long term consequences become more serious (higher c) or as adults' paternalism increases (higher δ). The role of λ will be the focus of the rest of the paper.

It is again instructive to compare the social learning case to the case where youngsters are informed and there are no signaling concern. In this case, since adults' behavior does not directly affect their children, Δp is zero and the cutoff is $\theta^*_I = -\overline{\boldsymbol{u}}$. It follows that

Proposition 2. *Whenever adults are paternalistic ($\delta > 0$) and have positive influence over their children ($\lambda > 0$), the propensity to abstain is higher under social learning than under fully informed youngsters. The distortion due to signaling is $\theta^*_{SL} - \theta^*_I = \frac{\lambda}{1+\lambda}\delta c > 0$.*

Proof. The result follows immediately by comparing θ^*_{SL} and θ^*_I.

Hence, whenever a degree of paternalism and of parental influence are both present, a distortion due to signaling emerges in equilibrium. Adults have ceteris paribus a lower propensity to choose self-gratification because they take into account the example they set to their children. Clearly enough, if whatever the parent craves is also deemed good for the child ($\delta = 0$), adults have no incentive to distort their behavior. A signaling distortion only arises when there is a misalignment of incentives. In this case, a parent will try to strategically manipulate her child's behavior by altering her own behavior. Similarly, if parents have no influence over their children ($\lambda = 0$), there is no point in teaching by example.

In Section 5, we use the two stage model to analyze the determinants of the signaling distortion. We argue that the comparative statics on λ generate empirical predictions that allow to test for the role of social learning. The next section looks instead at an OLG version of the model to analyze the long run distribution of behavioral traits in the population.

4 The long run distribution of behavioral traits

In this section, we extend the baseline model to consider long run dynamics. In a series of contributions, Bisin and Verdier (1998, 2001a, 2001b) analyze the mechanism generating the long term persistence of cultural/behavioral heterogeneity observed in modern societies. A common theme in these works is that convergence to a monomorphic population where all share the same behavioral trait can be avoided in the presence of *cultural sustitutability*, i.e. when the shrinking in popularity of a particular trait is matched by an increase in the parental influence of adults displaying that trait relative to other adults. Here, we ask what conditions are compatible with the long run persistence of heterogeneity when the effects of social learning are taken into account.

We consider an OLG setting where individuals live for two periods. In each period, agents play the two stage game by choosing behavioral trait $a \in \{0, 1\}$. Similar to the baseline model, when member Ai of the adult generation chooses $a \in \{0, 1\}$, he obtains a direct utility au_{Ai}, with $u_{Ai} = \overline{\boldsymbol{u}} + \theta_{Ai}$. Similarly, member Yi of the young generation choosing behavioral trait a obtains au_{Yi}. Note that we are not imposing pressure to conform at the outset. As will become clear, strategic interaction emerges *endogenously* in this version of the model. We follow existing literature in assuming

1. *Oblique/vertical transmission of preferences*: Youngsters inherit the taste of their role model.

2. *Imperfect Empathy:* Parents evaluate their children's utility using their own preferences

(augmented by paternalism).

These assumptions are not crucial but simplify the analysis and make the model comparable with canonical models of cultural transmission. The second assumption ensures that adult i's assessment of her child's (youngster i) utility is $a(u_{Ai} - \delta c)$ independently of whether vertical or oblique transmission occurs.

The information structure is similar to the baseline model. We assume that the support of θ is sufficiently large, so that youngsters, who do not know their taste shock, always mimic the behavioral trait of their role model. Upon becoming adults, agents learn their taste.[17] Different from the baseline model in which there was no past to be observed, adults can now also observe the frequency of each behavioral trait in the previous period (when they were youngsters). This typically allows them to extrapolate the value of θ from the size of heterogeneity in the recent past.[18] It will thus make the exposition less convoluted to simply assume that, upon reaching adulthood, agents can observe θ.[19]

The main difference with the previous sections is that we now endogenize the parental influence parameter. In line with the cultural transmission literature, we allow the probability that an adult adopting behavioral trait a becomes a role model for his child to depend on: 1) the behavioral trait she adopts and 2) the current prevalence of that trait in the population. We do not explicitly model how parental influence is determined but instead take a reduced form approach. Formally, let $\lambda_a : [0, 1] \to [0, 1]$ be a differentiable function capturing parental influence for trait $a \in \{0, 1\}$. We assume that the influence associated with each trait reflects the state of society at the time when adults choose whether or not to adopt a different trait

[17]While it is clear that agents who engage in self-gratification when young would learn their taste for it, it is less clear that agents who do not adopt the trait would necessarily learn it. We could assume that these agents only learn their taste with some positive probability. In this setting, agent who do not learn their tastes would simply keep mimicking the behavior of their role model in their adult life. This would bias transmission in favor of abstinence without affecting the qualitative results.

[18]Whenever the distance between θ and the adults' cutoff is less than ϵ, the share of adults choosing $a = 1$ is monotone in θ, so that observing the first involves learning the second.

[19]This has two effects. First, different from the baseline model where θ was not common knowledge, the game is not a global game and thus the equilibrium will not be generally unique. Second, since adults now know the average taste in the population, they could use this information to predict the taste of their child. This implies that, in general, $\mathbb{E}[u_{Yi}|u_{Ai}] \neq u_{Ai}$. Note, however, that, in order to avoid unnecessary complications, we have imposed imperfect empathy at the outset in this version of the model.

from that of their youth. As a result, the value of λ_a at time $t+1$ depends on the share, x_{Yt}, of time $t+1$ adults who have chosen $a=1$ in their youth. This captures the notion that, while adults can change their behavioral trait in the short term, the strength of vertical transmission adapts slowly to changes in society. For instance, people can quit smoking in days if needed, but it takes years for a child to form his personality. The underlying idea is that the lifestyle models proposed by society tend to reflect, rather than anticipate, the behavioral changes in the population. We should also note here that parental influence in our model should not necessarily be interpreted as the outcome of intentional socialization effort by parents, as in Bisin and Verdier (2001b). Rather, in the tradition of the socio-biology literature, the shape of λ_a simply reflects the presence of a conformist or anti-conformist bias in society. This means that some types of behavior may become more or less salient depending on the extent to which they conform to the dominant behavior in society.

4.1 Analysis

Since in equilibrium children conform to the behavioral trait of their role model, the total payoff of adult i adopting trait $a=1$ at time $t+1$ is

$$u_{Ai} + [\lambda_1(x_{Yt}) + (1-\lambda_1(x_{Yt}))x_{At+1}][u_{Ai} - \delta c], \tag{6}$$

where the first term is the adult's direct utility. The other terms are the adult's perception of the child's utility from $a=1$, given imperfect empathy and paternalism, weighted by the probability that the child will choose $a=1$ either through direct socialization, $\lambda_1(x_{Yt})$, or via oblique socialization, $(1-\lambda_1(x_{Yt}))x_{At+1}$. The equivalent expression when adult i adopts trait $a=0$ is

$$(1-\lambda_0(x_{Yt}))x_{At+1}[u_{Ai} - \delta c], \tag{7}$$

so that, using $u_{Ai} = \overline{\boldsymbol{u}} + \theta_{Ai}$, adult Ai's net payoff is[20]

$$\overline{\boldsymbol{u}} + \theta_{Ai} + [\lambda_1(x_{Yt}) + x_{At+1}(\lambda_0(x_{Yt}) - \lambda_1(x_{Yt}))][\overline{\boldsymbol{u}} + \theta_{Ai} - \delta c]. \tag{8}$$

This clearly reduces to the baseline model when $\lambda_0(x) = \lambda_1(x) = \lambda$, for all x. Since period t adults know (or can extrapolate) the value of θ, they can perfectly anticipate the equilibrium

[20] Note that strategic complementarity/substitutability emerge endogenously in this version of the model since the adult's net payoff depends on the population frequency x_{At+1}.

value, x^*_{At+1}, of x_{At+1}. They thus best respond by adopting a threshold strategy with cutoff θ^*_t, given by

$$\theta^*_{t+1} = \frac{\lambda_1(x_{Yt}) + x^*_{At+1}(\lambda_0(x_{Yt}) - \lambda_1(x_{Yt}))}{1 + \lambda_1(x_{Yt}) + x^*_{At+1}(\lambda_0(x_{Yt}) - \lambda_1(x_{Yt}))}\delta c - \overline{\boldsymbol{u}}. \tag{9}$$

Given an equilibrium share x^*_{At+1} of adults choosing $a = 1$ in period $t + 1$, the share of youngsters doing the same is

$$x^*_{Yt+1} = x^*_{At+1}[\lambda_1(x^*_{Yt}) + x^*_{At+1}(1 - \lambda_1(x^*_{Yt})) + (1 - x^*_{At+1})(1 - \lambda_0(x^*_{Yt}))]. \tag{10}$$

In words, a youngster can end up adopting self-gratification either through vertical transmission ($x_A\lambda_1$) or through oblique transmission ($x_A[x_A(1 - \lambda_1) + (1 - x_A)(1 - \lambda_0)]$). Since youngsters inherit the tastes of their role models, the distribution of tastes in society, $f^T_t : [\theta - \epsilon, \theta + \epsilon] \rightarrow \mathbb{R}^+$, evolves accordingly to the following law of motion,

$$f^A_{t+1}(z) = f^Y_t(z) = \begin{cases} f^A_t(z)[\lambda_1(x^*_{Yt}) + x^*_{At+1}(1 - \lambda_1(x^*_{Yt})) + (1 - x^*_{At+1})(1 - \lambda_0(x^*_{Yt}))] & z > \theta^*_t \\ f^A_t(z)[\lambda_0(x^*_{Yt}) + x^*_{At+1}(1 - \lambda_1(x^*_{Yt})) + (1 - x^*_{At+1})(1 - \lambda_0(x^*_{Yt}))] & z < \theta^*_t. \end{cases} \tag{11}$$

A *stationary state* is a tuple $(\theta^s, x^s_A, x^s_Y, f^s_A, f^s_Y)$, such that, for all t, $x_{At+1} = x_{At} \equiv x^s_A$, $x_{Yt+1} = x_{Yt} \equiv x^s_Y$, $f^A_{t+1} = f^A_t \equiv f^s_A$, $f^Y_{t+1} = f^Y_t \equiv f^s_Y$ for all z, and $\theta^*_{t+1} = \theta^*_t \equiv \theta^s$. Clearly, from (11), $f^s_A = f^s_Y \equiv f^s$. Given that each youngster imitates his role model, this in turn implies $x^s_A = x^s_Y \equiv x^s$. A stationary state must thus satisfy,

$$\lambda_0(x^s) = \lambda_1(x^s) \text{ if } x^s \notin \{0,1\} \tag{12}$$

$$\theta^s = \frac{\lambda_1(x^s) + x^s(\lambda_0(x^s) - \lambda_1(x^s))}{1 + \lambda_1(x^s) + x^s(\lambda_0(x^s) - \lambda_1(x^s))}\delta c - \overline{\boldsymbol{u}} \tag{13}$$

$$x^s = \begin{cases} 1 & \theta^s < \theta + \epsilon \\ \int_{\theta^s}^{\theta+\epsilon} f^s(z)dz & \theta^s \in [\theta - \epsilon, \theta + \epsilon] \\ 0 & \theta^s > \theta + \epsilon \end{cases} \tag{14}$$

The first condition comes from (10) once we impose the stationary requirement $x^*_{At} = x^*_{Yt}$. The existence of an interior stationary state thus requires that the functions λ_1 and λ_0 cross at least once, i.e. there exists $x^s \in (0, 1)$ such that $\lambda_0(x^s) = \lambda_1(x^s) \equiv \lambda^s$. Otherwise, stationary states necessarily involve homogeneous (i.e. monomorphic) populations, i.e. $x^s \in \{0, 1\}$. The second condition comes from the cutoff expression (9). The third condition is a simple consistency condition requiring that, given cutoff θ^s, the long term distribution of tastes is consistent with a share x^s of the adult population above the cutoff.[21]

The dynamics of the model are determined by the co-evolution of the population frequencies of behavioral traits and the distribution of tastes f_t^A. To gather intuition, suppose that, given x_{Yt}, $\lambda_1 > \lambda_0$ at $t+1$. This implies that individuals with taste terms above the cutoff θ^*_{t+1} are more likely to transmit their tastes to their children than individuals below the cutoff. In turn, this means that types above the cutoff will be overrepresented in the next generation relative to the the types below the cutoff. Clearly, this effect feeds back into the population frequencies, thus affecting λ_0 and λ_1 in ways that generate non-trivial dynamics.

Focusing on interior stationary states, conditions (12) and (13) together imply that the long run cutoff satisfies the familiar expression

$$\theta^s = \frac{\lambda^s}{1+\lambda^s}\delta c - \overline{\boldsymbol{u}} \tag{15}$$

where $\lambda^s \equiv \lambda_0(x^s) = \lambda_1(x^s)$ is now the (endogenous) stationary level of parental influence. To better illustrate how social learning affects the long run dynamics relative to standard models of cultural transmission, we focus on two canonical examples.

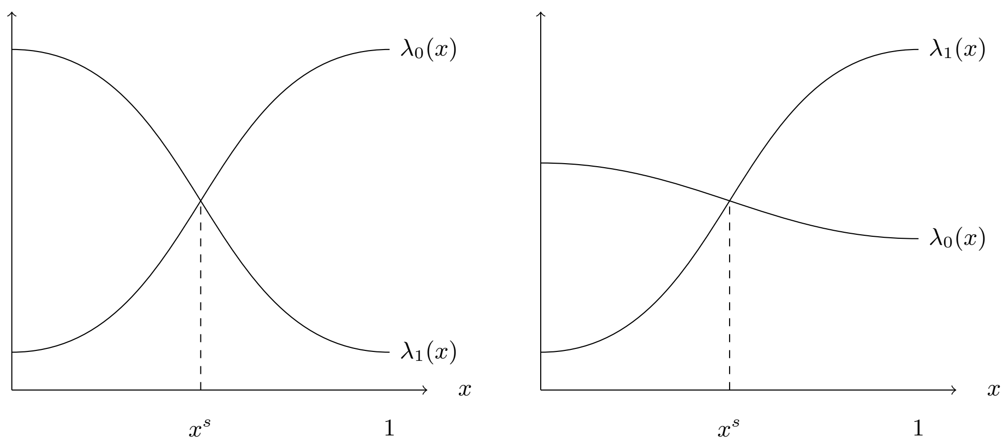

Figure 1: Parental influence with cultural substitutability and complementarity.

In the left panel of figure 1, λ_0 and λ_1 are symmetric, with λ_1 decreasing and λ_0 increasing in x. This is the textbook case of cultural substitutability: as the share of adults adopting trait $a = 1$ increases, the influence of parents choosing $a = 0$ increases while that of parents

[21]Note, however, that there will generally be multiple pairs (f^s, θ^s) yielding the same value of x^s in (13). Hence, while x^s is pinned down by the condition $\lambda_0(x^s) = \lambda_1(x^s)$, θ^s and f^s are not.

choosing $a = 1$ fades. In contrast, the right panel shows a situation where, as x increases, those adopting $a = 1$ acquire more influence relative to those choosing $a = 0$. In this case, there is clearly cultural complementarity (or a *conformist bias*). For instance, a conformist bias in transmission arises when parental smoking becomes more salient when most adults also smoke.

In the absence of signaling, the adults' propensity to choose either action would be exogenous with respect to x. In this case, the condition $\lambda_0(x) = \lambda_1(x)$ would pin down a stationary share x^s of agents choosing $a = 1$. This stationary share would be stable in the left panel but not in the right one. Intuitively, in the right panel, a small perturbation of x to the right of x^s would increase the influence of parents choosing $a = 1$ more than that of parents choosing $a = 0$. In turn, keeping adults' propensity to adopt trait $a = 1$ constant, this would increase the share of adults choosing $a = 1$ in the next period, moving society further away from the stationary state. This implies that, in the right panel, any stable stationary state must involve a monomorphic population.

Things change if we take into account signaling concerns. It is easy to check that the cutoff is increasing in x around x^s provided that

$$x\frac{d\lambda_0(x)}{dx} + (1-x)\frac{d\lambda_1(x)}{dx}\bigg|_{x=x^s} > 0. \tag{16}$$

In the right panel, this clearly holds if λ_0 is flatter than λ_1 at $x = x^s$. If the cutoff is locally increasing in x, the system need not converge to a monomorphic population. Intuitively, as in standard models, a small shock moving x to the right of x^s increases λ_1 relative to λ_0. Crucially, however, behavioral trait $a = 1$ is now associated with higher parental influence. This increases the signaling concerns of parents choosing $a = 1$, inducing some of them to switch to $a = 0$. This effect in turn reduces x and might potentially move society to the left of the stationary state, namely to a region where $a = 0$ carries larger influence. Explicitly accounting for social learning might thus generate very rich dynamics.

A full blown analysis of the transitional dynamics is technically complex.[22] We thus provide a simple illustration of how signaling may affect the local dynamics in the proximity of an interior stationary state.

[22] A prominent problem is that the long run distribution f^s will not be generally differentiable with respect to θ^s.

In general, the relationship between x_{Yt} and x_{At+1} is given by

$$x_{At+1} = \int_{\theta^*_{t+1}(x_{Yt})}^{\theta+\epsilon} f^A_{t+1}(z)dz, \tag{17}$$

whenever $\theta \in [\theta^*_{t+1} - \epsilon, \theta^*_{t+1} + \epsilon]$. Consider then the case where, close to the stationary state, both λ_0 and λ_1 are small, so that most of the transmission is oblique. Small values for λ_0 and λ_1 imply that the change in f^A_t from one period to the next is negligible (see equation 11). In the proximity of the stationary state (where $\lambda_0 = \lambda_1$), f^A_t can thus be reasonably approximated by the stationary distribution f^s. By the same token, when both λ_0 and λ_1 are small and close to each other, $x_{Yt} \approx x_{At}$. Equation (17) can thus be approximated by

$$x_{At+1} = \int_{\theta^*_{t+1}(x_{At})}^{\theta+\epsilon} f^s(z)dz. \tag{18}$$

Note that, while imposing small λs slows down the evolution of tastes, it does not affect qualitatively the way in which signaling concerns respond to changes in x_{At}. This is because the sign of the slope of the cutoff θ^*_{t+1} with respect to x_{At} depends on the slopes of λ_0 and λ_1, rather than their levels.[23]

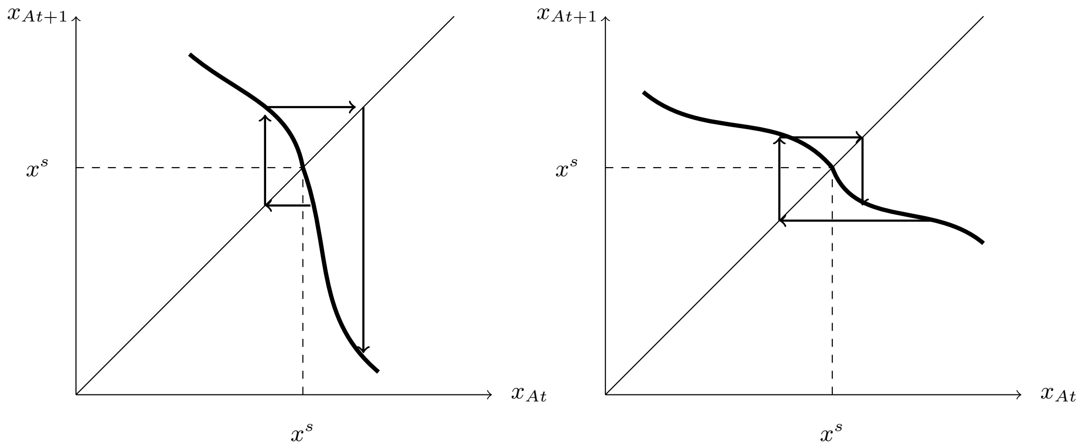

Figure 2: Dynamics around the stationary state.

Figure 2 provides two examples for equation (18). In both examples, the cutoff is increasing in the proximity of the stationary state, so that x_{At+1} and x_{At} are inversely related. Periods of higher than average frequency of $a = 1$ are thus followed by periods of lower than average frequency. Intuitively, an above average degree of self-gratification in one period

[23]Assuming small λs further simplifies the analysis by ensuring that the two stage game played in each period has a unique Perfect Bayesian Equilibrium.

raises the parental influence associated with self-gratification, so that some adults switch to abstinence. In turn, this lowers the influence of self-gratification in the next period inducing some to switch to self-gratification. In the left panel, small changes in x_A yield sharp movements of the cutoff. This happens if x^s is not too large and λ_1 is sufficiently steep relative to λ_0 in the proximity of x^s. As shown in the figure, in this case society will oscillate away from the stationary state over time. In contrast, in the right panel, the cutoff is less sensitive, so that a cycle converging to the interior stationary state is possible.

As noted above, a positive slope for the cutoff is not inconsistent with cultural complementarity. Hence, the cyclical behavior around the stationary state can emerge in spite of a conformist bias in cultural transmission. Finally, the assumption that the λs are small close to the stationary state helps to illustrate the role of signaling in a stark way by slowing down the evolution of preferences. Similar results would obtain without this assumption so long as the effect of signaling is sufficiently strong to offset the drift toward a monomorphic equilibrium induced by the transmission of preferences.

To sum up, when signaling concerns and paternalism are taken into account, changes in the share of adults choosing self-gratification affect parental influence. Through the signaling mechanism illustrated in Section 3, this may change the propensity to choose self-gratification in a way that allows heterogeneity to persist even in the presence of a conformist bias.

5 Empirical implications

In this section, we return to the baseline model to derive testable implications. The two stage model – where λ is exogenous – is well suited to an empirical investigation on household level data. The co-determination of (trait dependent) influence and population frequencies of behavior analyzed in the previous section is to a large extent a dynamic phenomenon, thus difficult to detect with cross-sectional data. Furthermore, we do not have empirical measures of trait-specific parental influence, but only measures of generic influence. In other words, the data tell us the extent to which the parent sees herself as a role model, but not her specific influence in relation to smoking, drinking, etc. Any variation across adults in perceived parental influence is thus unlikely to be driven by the population frequencies of any specific behavioral trait. That said, in order to avoid a causal interpretation of our empirical results, we cast the model's predictions in terms of simple correlations.

The central prediction of the theory is contained in the following Corollary to Proposition 2,

Corollary 2. *Whenever $\delta c > 0$, the signaling distortion is increasing in the degree of parental influence, λ. Under social learning, parental influence and the adults' propensity to choose self-gratification are thus inversely related. This effect is not present in the absence of signaling concerns.*

The model has also implications for children behavior. Clearly enough, since children imitate their role model, we expect some correlation between the behavior of the parent and that of the child. However, this prediction is not unique to social learning as positive correlation between parent and child would emerge in virtually any model of cultural transmission. More relevant for the mechanism considered here is the fact that parental influence has an indirect effect on the observed behavior of children. Higher parental influence increases the chances that adult Ai chooses abstinence and thus, through vertical transmission, reduces the likelihood that his child will engage in self-gratification.

Corollary 3. *Whenever $\delta c > 0$, the likelihood that a youngster will engage in self-gratification is inversely related to parental influence.*

The relationship between influence and the signaling distortion is deceptively simple. In order to gain a better understanding of the effects at work, we need to introduce in the model some heterogeneity in parental influence.

5.1 Parental differences in influence and the role of conformity

The empirical predictions of the baseline model have been cast in terms of comparative statics for a society where adults share the same degree of parental influence. We now extend the model to allow for heterogeneity in influence. There are two main reasons for analyzing this case. First, most of the evidence available (and the one presented in the Section 6) comes from differences in parental influence among members of the same society. For practical purposes, we thus want to recast Corollaries 2 and 3 in terms of correlation between parental influence and behavior at individual level. The second reason is that considering heterogeneity in parental influence helps to understand the role of conformity and how it complicates the task of empirically measuring the distortion due to social learning.

We thus assume that there are two groups of adults in the population: High influence adults (with influence parameter λ_H) and low influence adults (with parameter $\lambda_L < \lambda_H$). The fraction of high influence adults is denoted with $\mu \in (0, 1)$. We assume that the existence of the groups (and the parameters λ_H, λ_L, and μ) is common knowledge. Moreover, each youngster knows whether his role model is in the high influence or in the low influence group. A possible (but not necessary) interpretation for these assumptions sees low influence adults as adults who are childless or geographically separated from their children. Another interpretation is that these adults simply have less chances to interact with their children. In experimental settings where the experimenter can control how salient an adult's behavior is to her child (Ben-Ner et al. 2015, Houser et al. 2016), type H adults can be interpreted as a "treatment" group of adults who are fully visible to their children while type L adults are the control group.

It will help the discussion to reduce the strength of the pressure to conform to a single parameter. This is obtained by assuming that the v function has symmetric upper and lower bounds: i.e. $v(1, 1) = -v(0, 0) \equiv \overline{v} > 0$. In other words, the extra utility from choosing self-gratification when everyone else does the same compared to a situation where everyone chooses abstinence is $2\overline{v}$. The smaller the value of $\overline{v}$, the flatter is the v function (and thus the weaker the conformity motive). The next result characterizes the equilibrium in the model with heterogeneous influence

Proposition 3. *For D sufficiently large there is a unique threshold equilibrium and it is such that: 1) Each youngster chooses self-gratification if and only if he observes his role model choosing self-gratification; 2) High and low influence adults choose self-gratification if their taste shock θ_{Ai} is above thresholds θ^*_H and $\theta^*_L < \theta^*_H$, respectively, where*

$$\theta^*_H \geq \frac{\lambda_H}{1 + \lambda_H}\delta c - \overline{\boldsymbol{u}} - (1 - \mu)\overline{v} \tag{19}$$

$$\theta^*_L \leq \frac{\lambda_L}{1 + \lambda_L}\delta c - \overline{\boldsymbol{u}} + \mu\overline{v} \tag{20}$$

with the above expressions holding with equality whenever

$$\delta c \frac{\lambda_H - \lambda_L}{(1 + \lambda_H)(1 + \lambda_L)} > \overline{v} + 2\epsilon \tag{21}$$

Proof. See supplementary material.

The RHS of (19) and (20) provide explicit solutions for the two groups' propensities to abstain when inequality (21) holds. When this condition does not hold, the expressions provide a lower and upper bound for θ_H^* and θ_L^*, respectively. Qualitatively, the same arguments apply independently of whether (21) holds or not, but the analysis is slightly more convoluted in the latter case.

Unsurprisingly, adults with high influence tend to have a lower propensity to choose self-gratification than low influence adults (i.e. $\theta_H^* > \theta_L^*$). In the absence of signaling concerns, the two groups would obviously have identical propensity. If one can measure parental influence at the individual level – or can manipulate it in an experimental setting – it is then possible to test Corollaries 2 and 3 using variation in parental influence across individuals.

That said, Proposition 3 also illustrates the central difficulty in using individual differences in influence. Consider the simple case where (21) holds. It is immediate to verify that, as the pressure to conform increases (i.e. a larger $\overline{v}$), θ_H^* becomes smaller whereas θ_L^* becomes larger. In other words, the need to coordinate makes each group's propensity to abstain closer to the one of the other group. This is not only driven by the fact that, in the presence of complementarity, adults want to conform to the rest of society. They also anticipate that their children will similarly want to conform. Whether it is the high influence adults who mainly conform to the low influence ones or the opposite depends on the relative size of the two groups (μ).

It is clear that pressure to conform, by inducing more coordination, generally makes the task of empirically detecting the effect of social learning using variation in parental influence more complicated. One can in principle identify two extreme classes of behavioral traits

1. Traits with a strong individualistic connotation where the pressure to conform is small or non-existent (e.g. whether to take coffee with sugar or without)

2. Traits with a strong social connotation where the pressure to conform is strong (e.g. participation in religious rituals).

In the first case, individual differences in parental influence will translate into relatively large differences in the propensity for self restrain. In the second case, the propensity for self restrain is mostly shaped by the average parental influence in the society and individual differences are less important. That said, most activities of interest are not easily classified in either of the two categories above. For instance, eating, smoking, and drinking are in part social activities. Similarly, work habits reflect individual preferences but also conformism. It is clearly possible to design experiments that isolate the effect of parental influence from the confounding effects of other social interactions, as for instance in Ben-Ner et al. (2015) or Houser et al. (2016). However, whether this effect is strong enough to explain differences among individual behavior in situations where the pressure to conform is non-negligible remains an open question. In Section 6, we provide some evidence in this direction.

6 Evidence on parental influence and parental self-restraint

In this section we examine two empirical implications unique to the social learning model. The first implication is stated in Corollary 2: parental influence is inversely related to parental self-gratification. To do this we use micro-data with the novel property of containing information, for mothers and fathers, on (perceived) parental influence. We interrogate the correlation between parental influence and two behaviors which fit well with the model, smoking and excessive drinking. The second implication is stated in Corollary 3: parental influence is inversely related to child self-gratification. Here we focus the analysis on the child's propensity to try smoking. As a secondary empirical exercise we use these data to examine evidence of the intergenerational transmission of behavior from parents to children.

We emphasize that the purpose of this exercises is to test two qualitative predictions of the social learning model. Estimation of structural parameters underlying the the social learning model, while interesting, is beyond the scope of this paper.

6.1 Data

Data are drawn from wave A of the Iowa Youth and Families Project (Cogner et al., 2011). The Iowa Youth and Families Project was designed to study how families in rural Iowa coped with the economic stress that resulted from the 1980s farm crisis. The study collected detailed household information from interviews, conducted in 1989, with 451 households containing a child in seventh grade (12–14 year old). This child is referred to as the *reference child*. Data include information on the reference child, the mother and the father.

For our purpose, these data have the unique property of containing information, for each of the mother and the father, on beliefs about the influence that parents have on their children's behavior. We create an index of *perceived parental influence* based on the five questions presented in Table 1 (mother and father question labels in parenthesis). For each question, mothers and fathers are separately asked to report the extent to which they agree with each statement on a 1–5 scale (1 = strongly agree, 5 = strongly disagree).

Table 1: Questions underlying the perceived parental influence index

a)	The behavior of parents largely determines a child's self-concept (am202024, af202024).
b)	Parents don't have much impact upon what their children grow up to be like (am202025, af202025).
c)	Kids grow up to be a lot like their parents (am202026, af202026).
d)	When kids are getting into trouble or having problems it is usually related to the behavior of their parents (am202027, af202027).
e)	Parents have a very strong influence upon the values of their kids (am202028, af202028).

Notes: Question labels for mother and father in parenthesis.

In question b), a higher value reflects a greater agreement with the influence of parents, while in the remaining questions a higher value reflects greater disagreement with the influence of parents. We compute an index of parental influence by calculating the difference $b - a - c - d - e$. An increase in the index value reflects a greater agreement that parental influence is important. Observations are binned into one of three groups reflecting *high*, *medium* and *low* beliefs, according to three quantiles of the index value.

These data also contain information that allows us to consider the extent to which a child is "vertically socialised". For each child we consider the answer to the following question (asked separately for reference to the mother and father): "When I grow up, I'd like to be like my father[mother]". Children report the extent to which they agree with each statement on a 1–5 scale (1 = strongly agree, 5 = strongly disagree). We use this information to create a single dummy variable equal to 1 if the child *strongly agrees* or *agrees* with respect to the question framed to either the mother or the father, and 0 if the child answers otherwise. We interpret this as a measure of the degree of vertical socialization; children who do not wish to be like either parent are more likely influenced by external role models.

We focus the analysis on two behaviours relevant to the social learning model as presented in this paper. For each of the mother and father we look at whether the parent is a smoker or non-smoker and how many times in the past month the parent consumed four or more alcoholic drinks in a single day. We also look at a different behavior, likely not viewed as setting a bad example, how frequently do parents attend religious ceremonies[24]. In addition, the analysis makes use of a number of control variables including household income, age and education for mother and father, the reference child's sex, age, and the number of older and younger siblings. Summary statistics are reported in Table 2, with further summary statistics and the details of relevant variables provided in the supplementary appendix to this paper.

Mean values of each variable are reported for low, medium or high perceived influence in columns 1, 2 and 3 of Table 2. Variables which refer to child or household characteristics are reported by father's perceived influence. Column 4 presents p-values corresponding to an F-test of the null hypothesis that the mean values are the same across all parental influence quantiles. In the case of education categories the p-value corresponds to a chi-squared test of equality in the distribution across each education category.

A number of details are worth noting. First, smoking and drinking are both lower for parents with higher perceived influence (although only smoking is significantly different across the quantiles) and church attendance is increasing with the perceived influence index. Education, for both mother and father, significantly differs across the quantiles; parents with a higher perceived influence also tend to be more educated. Household earnings are also increasing with perceived influence, although the difference is not statistically significant. Finally, mean values for almost all other characteristics are statistically indistinguishable across the three quantiles (the exception being the number of siblings older than the reference child). Interestingly, this also applies to our measure of vertical socialization: whether or not a child reports wishing to be like at least one of their parents.

6.2 Parental influence and parental behavior

The simple correlation between parental smoking/drinking and perceived parental influence reported in Table 2 is consistent with the hypothesis of the social learning model: parents who believe they have a greater influence on their children are less likely to engage in these

[24]Based on a derived frequency. See supplementary appendix for details.

Table 2: Descriptive statistics

		Perceived parental influence			
		(1)	(2)	(3)	(4)
Father		*Low*	*Medium*	*High*	*p-value*
	Smoker	0.455	0.412	0.256	0.002
	Drinks (days per month)	1.935	1.447	1.224	0.249
	Church (per year)	34.735	36.929	45.600	0.020
	Age	40.340	39.300	39.536	0.140
	Education				
	< High school	0.077	0.024	0.024	
	High school	0.545	0.376	0.376	
	PS, < 4 years	0.269	0.353	0.248	
	4-year degree	0.083	0.165	0.256	
	> 4-year degree	0.026	0.082	0.096	0.000
Mother					
	Smoker	0.319	0.164	0.212	0.005
	Drinks (days per month)	0.748	0.503	0.219	0.180
	Church (per year)	39.378	47.624	47.669	0.047
	Age	37.519	37.879	37.662	0.747
	Education				
	< High school	0.044	0.012	0.013	
	High school	0.541	0.436	0.331	
	PS, < 4 years	0.311	0.375	0.411	
	4-year degree	0.096	0.152	0.219	
	> 4-year degree	0.007	0.018	0.026	0.009
Child†					
	Wish to be like parent?	0.558	0.536	0.584	0.708
	Ever tried smoking?	0.205	0.271	0.136	0.019
	At least one parent smokes	0.564	0.512	0.288	0.000
	Age	12.577	12.624	12.648	0.532
	Sex (male=1)	0.474	0.488	0.464	0.917
	Only child	0.308	0.376	0.352	0.423
	Number of younger siblings	0.603	0.582	0.704	0.382
	Number of older siblings	0.391	0.335	0.200	0.027
	Household earnings	24,263	27,223	28,047	0.127
	Observations	156	170	125	

Notes. *p-value* reflects a test of the null hypothesis that mean values are equal for low, medium and high perceived parental influence. †Perceived parental influence for child's variables reflects that of the father.

behaviors. However, this relationship is also open to alternative explanations. For example, there is a well documented negative relationship between education and smoking (see for example, de Walque, 2007). In Table 2 we see that parents with higher levels of eduction perceive a greater influence of parental behavior on children's behavior. This leads to concern that the correlation between smoking/drinking and parental influence may be spurious, attributable to differences in underlying preferences (reflected in education choices). For example, relative to parents with high future discount rates, it is likely that parents who discount the future less heavily receive more education, smoke and drink less, and place more weight on their children's long run welfare. These parents therefore have an incentive to believe they have more influence on their children. If this explanation accurate, conditioning on education will significantly decrease the smoking/drinking and perceived parental influence correlation observed in Table 2.

To account for this we consider the following regression:

$$y_i^p = \alpha_0 + \alpha_M^p PPIM_i^p + \alpha_L^p PPIL_i^p + AGE_i^{p\prime}\Gamma_1 + EDUC_i^{p\prime}\Gamma_2 + X_i'\Gamma_3 + \zeta_i^p \tag{22}$$

where y_i^p is the outcome of interest (smoking, drinking, church) for parent $p \in$ {mother, father} in household i. $PPIM_i^p$ and $PPIL_i^p$ are dummy variables to indicate the parent belonging to either the medium or low perceived parental influence quantile. AGE_i^p is a vector which includes a quadratic term for parent p's age, and $EDUC_i^p$ is a vector of education dummy variables corresponding to parent p to allow for a discontinuous impact of education. X_i is a vector of household and child variables, including sex and age of the reference child, an only-child indicator, separate variables for the number of older and younger siblings, and annual household earnings[25]. ζ_i^p reflects the influence of unobserved variables. The coefficients of interest, α_M^p and α_L^p, reflect the incremental effect of parent p having a medium or low perceived influence relative to parent p having a high perceived influence. This interpretation of the estimated values $\hat{\alpha}_M^p$ and $\hat{\alpha}_L^p$ is subject to the assumption that any unobserved variables that influence the outcome are uncorrelated with perceived parental influence.

We report different estimates of α_M^p and α_L^p in Table 3. In panel A we report a regression of each outcome on the medium and low quantile dummies, excluding control variables.

[25] Six observations are omitted from the analysis due to missing information on household earnings.

To this we compare estimates for equation (22) using OLS (panel B) and propensity score matching[26] (panel C). The comparison of panel A to panels B and C provides information about the importance of unobserved characteristics, such as discount factors. If these unobservables are important then we expect to see large changes in the estimated coefficients across the different estimates. Panels B and C provide two different methods for estimation based on matching on observables[27]. If estimates are robust, we expect these methods to be in agreement with one another.

The estimates reported in Table 3 are generally consistent with the predictions of the social learning model. Among fathers with a high perceived parental influence, 25.5% smoke. Smoking is 15.6 and 19.9 percentage points higher among fathers with indexes in the medium and low quantiles. Similarly, heavy drinking is more common among fathers with low and medium perceived influence than those with high perceived influence, although these estimates are noisy. As expected, church attendance does not behave in the same way. We see significantly greater church attendance among parents with high perceived influence. This is consistent with church attendance reflecting a behavior into which parents hope to positively influence their children.[28]

The non-trivial increase in the R^2, between panels A and B, suggests that the additional control variables play a significant role in explaining the outcomes. However, the patterns are similar to what is observed in the unconditional correlations. Controlling for parental education, age and a number of other household characteristics we find that fathers who have a low perceived influence index are between 16.7 and 18.1 percentage points more likely to smoke than are fathers who have a high perceived influence index, very close to what was observed in panel A. For mothers, a low perceived influence index is associated with between a 6.8 and 14.7 percentage point increase in smoking, relative to a high perceived influence index, with the propensity score estimates statistically significant. Similar estimates are found for drinking, which is statistically significant and positive for both mothers and

[26]See the supplementary material for the full set of reported coefficients and balancing tests for the propensity score matching.

[27]Angrist (1998) shows that OLS regression is equivalent to a matching method which weights according to the variation in treatment across different control values. Propensity score matches based on the probability that an observation receives treatment. In the current context, 'treatment' refers to either $PPIM_i^p = 1$ or $PPIL_i^p = 1$.

[28]See also Patacchini and Zenou (2016) for an analysis of the transmission of religiosity.

Table 3: Parent behavior and parent's perceived influence in child

	Father			Mother		
	Smoking	Drinking	Chruch	Smoking	Drinking	Chruch
A. Unconditional						
Medium	0.156***	0.223	-8.671**	-0.048	0.284	-0.0446
	(0.054)	(0.384)	(4.059)	(0.044)	(0.219)	(3.693)
Low	0.199***	0.711	-10.86***	0.107**	0.530*	-8.291**
	(0.056)	(0.451)	(4.118)	(0.052)	(0.279)	(3.797)
N	451	451	451	451	451	451
R^2	0.023	0.002	0.013	0.019	0.003	0.009
B. Conditional						
Medium	0.141***	0.300	-6.098	-0.061	0.339	1.453
	(0.055)	(0.393)	(3.981)	(0.044)	(0.255)	(3.566)
Low	0.167***	0.606	-6.460	0.068	0.533*	-4.655
	(0.060)	(0.485)	(4.114)	(0.054)	(0.286)	(3.846)
N	451	451	451	451	451	451
R^2	0.056	0.065	0.074	0.071	0.024	0.092
C. Propensity score						
Medium	0.149*	0.745	-4.025	-0.043	0.340	2.190
	(0.086)	(0.496)	(6.739)	(0.064)	(0.251)	(5.209)
Low	0.181**	1.138*	-0.362	0.147**	0.620**	-7.829
	(0.085)	(0.621)	(6.439)	(0.072)	(0.301)	(6.137)
N (medium)	280	280	280	308	308	308
N (low)	261	261	261	275	275	275
Mean values for high perceived influence	0.256	1.224	45.60	0.212	0.219	47.67

Notes. Robust standard errors reported in parenthesis. *, **, ***, denote statistical significance at 10%, 5% and 1%. Panel *A* and Panel *B* report the results of unconditioned and conditioned linear regressions of behavior on perceived influence (high influence is the excluded category). Conditioned regressions control for parent age (quadratic), parent eduction, age and sex of reference child, number of older and younger siblings, only child status. See data appendix for regression details as well as all estimated coefficients.

fathers in the propensity score estimates. For church attendance, all differences between high and low parental influence are negative, although not statistically significant.

The analysis we present here suggests that parents who perceive parental behavior to influence the behavior (or future behavior) of the child are less likely to engage in smoking and drinking and more likely to attend church regularly. We interpret this as being consistent with parents behaving strategically, in a manner that conforms with the social learning model. This is also consistent with the evidence from experimental studies. Houser et al. (2016) create a coin-toss experiment in which parents have the opportunity to cheat. Varying the presence of children during the experiment, they find that the presence of a daughter significantly reduces the tendency of parents to cheat. While the experimental setting allows them to cleanly estimate the effect of child presence on parental behavior, they cannot distiguish between two motives: does the parent not wish to appear dishonest to the child or, does the parent wish to transmit pro-social behaviour to the child? Ben-Ner et al. (2015) report that parents increase their contributions in the dictator game when they are aware that these will be shown to their children.

6.3 Parental influence on child behavior

The social learning model states that children will be influenced by the behavior of their parents. To explore this in our data we focus on parent and child smoking.[29] Intergenerational transmission suggests that we should expect to see a positive relationship between child smoking (or experimentation with smoking) and parental smoking. However, a positive correlation from an OLS regression is consistent with alternative mechanisms of intergenerational transmission and is not unique to the social learning model. Therefore, we also provide a test of Corollary 3: youngster smoking is inversely related to parental influence.

We start by estimating the relationship between child and parent smoking based on the following equation:

$$SMOKE_i^C = \beta_0 + \beta_1 SMOKE_i^P + H_i'\Pi + \mu_i, \tag{23}$$

where $SMOKE_i^C$ is a variable equal to 1 if the reference child in household i has ever smoked tobacco, and 0 otherwise. $SMOKE_i^P$ is a variable equal to 1 if either the mother or father

[29] The focus on smoking is due to the fact that this is the variable for which we both have reliable information on parent and child choices, and which yields relatively precise estimates in Table 3.

Table 4: Parental influence on child smoking

Outcome: Child smokes	*Parent smoking*			*Perceived parent influence*		
	OLS		*IV*		External role model	Parental role model
	A1	A2	B1	C1	C2	C3
Parent smokes	0.112***	0.097**	0.334**			
	(0.038)	(0.042)	(0.149)			
Perceived parent influence			*First stage*			
Low (father)			0.235***	0.048	-0.068	0.132**
			(0.060)	(0.048)	(0.084)	(0.056)
Medium (father)			0.223***	0.111**	-0.022	0.189***
			(0.055)	(0.046)	(0.080)	(0.056)
Low (mother)			0.124**	0.068	0.102	0.068
			(0.055)	(0.052)	(0.086)	(0.063)
Medium (mother)			-0.089*	0.004	0.027	-0.008
			(0.054)	(0.045)	(0.075)	(0.058)
Controls included	No	Yes	Yes	Yes	Yes	Yes
N	451	445	445	445	196	249
R^2	0.019	0.104	0.025	0.108	0.165	0.148
F-stat (first stage)			9.370			
Sargan's χ^2 (p-value)			0.376			

Notes. Robust standard errors reported in parenthesis. *, **, ***, denote statistical significance at 10%, 5% and 1%. Coefficients for separate regressions reported in each column. Columns A1 and A2 reflect OLS regression of child smoking on parent smoking without and with control variables. Column B1 reflects the IV regression of child smoking on parent smoking. Perceived influence index variables are the excluded instruments. First stage results reflect regression of parent influence on parent smoking. First stage F-stat corresponds to a test of the joint significance of all excluded instruments in the first stage regression. Sargan's χ^2 reflects a test of the null hypothesis that excluded instruments are exogenous. Columns C1, C2 and C3 regress child smoking on parental influence for all children and limiting to children with external role models and internal role models, respectively.

in household i smoke, and 0 otherwise. H_i is a vector of control variables reflecting the characteristics of parent, household, and child. H_i includes age and education level of both parents, sex and age of the reference child, an only-child indicator, separate variables for the number of older and younger siblings, and annual household earnings. μ_i reflects the influence of unobserved variables on $SMOKE_i^C$. The coefficient of interest, β_1, reflects the incremental probability that a child smokes when one (or both) of his or her parents smoke, conditional on H_i.

The regression results are reported in columns A1 (unconditional) and A2 (conditional on H_i) of Table 4. Coefficient estimates suggest that a child with a parent who smokes is about 10 percentage points more likely to smoke. While these results are consistent with the existence of intergenerational transmission, our theory implies that they are likely to exhibit a downward bias as a result of the endogenous relationship between parental and child smoking. If parents indeed behaved strategically, they would tend to avoid smoking precisely when they are concerned that their child might be experimenting with smoking. To account for this we estimate the relationship in equation (23) using parents' perceived influence as an instrument for parental smoking. This strategy is summarized in the following equations:

$$SMOKE_i^P = \eta_0 + \sum_p \left(\eta_M^p PPIM_i^p + \eta_L^p PPIL_i^p\right) + H_i'\Lambda + \omega_i \quad (24)$$

$$SMOKE_i^C = \beta_0 + \beta_1 \widehat{SMOKE}_i^P + H_i'\Pi + \mu_i. \quad (25)$$

All variables in the second equation are the same as those in equation (23), with the exception of $\widehat{SMOKE}_i^P$. $\widehat{SMOKE}_i^P$ is the predicted value from the first equation, in which parental smoking is regressed on dummies for medium and low perceived parental influence for both mother and father, and all control variables contained in H_i.

We approach the instrumental variable procedure with a note of caution. For perceived parental influence to be a valid instrument it must only impact child's smoking through parental smoking. This assumption is clearly violated if the parent's perceived influence changes another behavior and that behavior influences both child and parent smoking. For example, if parents with a high perceived index are more likely to attend church, and church attendance influences smoking, this would violate our assumption. With this in mind we cautiously interpret the instrumental variable estimates.

The coefficient reported in B1 suggests a strong parental effect. Parental smoking leads to a 33.4 percentage point increase in child smoking. Further, this results suggests that the magnitude of the intergenerational effect is underestimated by OLS, potentially due to parents of children with higher predisposition to smoke being more likely to pre-emptively forgo smoking. The coefficients for the parent's perceived index reflect the first stage regression of parental smoking, and are consistent with the results reported in Table 3. We test the over-identifying restrictions of the model using Sargan's χ^2, from which we fail to reject the null hypothesis that the excluded instruments are exogenous.[30]

Finally, we present a direct test of Corollary 3 from the social learning model: child smoking is inversely related to parental influence. To do this we estimate equation (23), replacing $SMOKE_i^P$ with a vector of perceived parental influence indicators for both parents, $\{PPIM_i^{father}, PPIL_i^{father}, PPIM_i^{mother}, PPIL_i^{mother}\}$. The corresponding coefficients, representing the incremental probability of a child smoking relative to having parents with a high influence index value, are reported in column C1. There is a weak, positive correlation between being in a lower quantile and child smoking (although this is only significant for fathers medium index values). In columns C2 and C3 we repeat this regression for the sub-groups of children who do not perceive their parents as role models and for those who do[31], respectively. As expected, we do not find a significant correlation between child smoking and parental index values when the child does not view parents as role models. However, when parents are role models, we find that the father's perceived influence is significantly correlated with child smoking. We interpret this result as consistent with Corollary 3. A child is 12.3 percentage points more likely to smoke if his or her parent is in the lowest index group (relative to the highest).

7 Concluding remarks

What is the object of cultural transmission? Does it consist in the transmission of preferences? Or does it involve the transmission of beliefs? Or perhaps a combination of both? The answers to these questions have far reaching consequences not just for our understand-

[30]However, it should be noted that the validity of this test is dependent on at least one of the excluded instruments being exogenous.

[31]Based on the "When I grow up, I'd like to be like my father[mother]" question discussed in section 6.1

ing of the world, but also for policy and welfare analysis. For instance, while economists typically acknowledge the fact that *de gustibus non est disputandum*, beliefs can be shown to be false or unjustified. In spite of this, the literature on cultural transmission has almost entirely focused on the transmission of preferences, largely neglecting the role of beliefs. This paper contributes to the literature by giving beliefs the center stage. Our model shows that the transmission of beliefs from adults to children has precise implications for the strategic behavior of parents. This approach thus provides an indirect way to assess the relevance of beliefs in the transmission of culture. Moreover, some of these effects are quite subtle and may inform the policy debate.

Empirically, we find some evidence for the prediction that parents strategically adopt behaviors aimed at setting examples to their children. The extent to which a parent sees herself as a role model is inversely related to her propensity to consume alcohol or tobacco and to the likelihood that her child will experiment with smoking.

Finally, and clearly enough, we should mention that our narrow focus leaves out important aspects of socialization, like for instance school, neighborhood, or marriage choices. Future work should therefore be aimed at incorporating these elements into our theory.

References

[1] Adriani, F., and Sonderegger, S., (2009) Why do parents socialize their children to behave pro-socially? An information based theory. *Journal of Public Economics*, 93: 119-1124.

[2] Adriani, F., and Sonderegger, S., (2016) Signaling about norms: Socialization under strategic uncertainty. *Scandinavian Journal of Economics*, forthcoming.

[3] Alesina, A., and Giuliano, P., (2014) Family Ties. *Handbook of Economic Growth*, edited by Aghion, P., and Durlauf, S.N. 2A: 177-215. The Netherlands: North Holland.

[4] Angrist, J., (1998) Estimating the labour market impact of voluntary military service using social security data on military applicants. *Econometrica*, 66: 249-288.

[5] Bandura, A., and Walters, R., (1963) *Social learning and personality development.* New York, Holt, Rinehart and Winston.

[6] Banerjee, A.V., (1992) A Simple Model of Herd Behavior. *Quarterly Journal of Economics*, 107: 797-817.

[7] Bénabou, R., and Tirole, J., (2006) Belief in a Just World and Redistributive Politics. *Quarterly Journal of Economics*, vol. 121, n. 2, p. 699-746.

[8] Ben-Ner, A., List J.R. , Putterman, L. and Samek, A., (2015) Learned Generosity? A Field Experiment with Parents and their Children. CESR-Schaeffer Working Paper.

[9] Bernheim, B.D., (1994) A theory of conformity. *Journal of Political Economy*, 102:841-877.

[10] Bikhchandani, S., Hirschleifer, D., and Welch, I., (1992), A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100(5), 992- 1026.

[11] Bisin, A., and Topa, G., (2003) Empirical Models of the Transmission of Culture: Some Examples. *Journal of the European Economic Association*, 1: 363-75.

[12] Bisin, A., and Verdier, T., (1998) On the cultural transmission of preferences for social status. *Journal of Public Economics* 70: 75-97.

[13] Bisin, A., and Verdier, T., (2001a) The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory*, 97:298-319.

[14] Bisin, A., and Verdier, T., (2001b) Agents with Imperfect Empathy May Survive Natural Selection. *Economics Letters*, 2:277-85.

[15] Bisin, A., and Verdier, T., (2011) The Economics of Cultural Transmission and Socialization, in *Handbook of Social Economics*, Benhabib, J., Bisin, A., and Jackson, M., eds. Elsevier.

[16] Boyd, R. and Richerson, P., (1985) *Culture and the Evolutionary Process*. Chicago, University of Chicago Press.

[17] Bricker, J., Otten, R., Liu, J., and Peterson, A. (2009), Parents who quit smoking and their adult children's smoking cessation: a 20-year follow-up study. *Addiction*, 104: 10361042.

[18] Calvo-Armengol, A., and Jackson, M.O., (2009) Like Father, Like Son: Social Network Externalities and Parent-Child Correlation in Behavior. *American Economic Journal: Microeconomics*, 1: 124-50.

[19] Carlsson, H., and Van Damme, E. (1993) Global games and equilibrium selection. *Econometrica* 61:989-1018.

[20] Cavalli-Sforza, L. and Feldman, M., (1981) *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton, Princeton University Press.

[21] Cervellati, M., and Vanin, P., (2013) Thou shalt not covet: Prohibitions, temptation and moral values. *Journal of Public Economics*, 103: 15-28.

[22] Conger, R.D., Lasley, P., Lorenz, F.O., Simons, R., Whitbeck, L.B., Elder, G.H. Jr., and Norem, R. Iowa Youth and Families Project, 1989-1992. ICPSR26721-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2011-11-03. https://doi.org/10.3886/ICPSR26721.v2

[23] Corneo, G., and Jeanne, O., (2009) A Theory of Tolerance. *Journal of Public Economics* 93:691-702.

[24] de Walque, D. (2007), Does education affect smoking behaviors? Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics* 26: 87795.

[25] Doepke, M., and Zilibotti, F., (2011) Intergenerational Transmission of Risk Preferences, Entrepreneurship, and Growth. Mimeo.

[26] Doepke, M., and Zilibotti, F., (2014) Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission. *Econometrica*, forthcoming.

[27] Fernandez, R., Fogli, A., and Olivetti, C., (2004) Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *Quarterly Journal of Economics*, 119: 1249-1299.

[28] Galea, S., Nandi, A., and Vlahov, D., (2004) The social epidemiology of substance use. *Epidemiologic Reviews* 26: 35-52.

[29] Houser, D., List, J.R., Piovesan, M., Samek, A. and Winter, J., (2016) Dishonesty: from parents to children. *European Economic Review* 82, 242-254.

[30] Jackson, C., and Henriksen, L., (1997) Do as I say: Parent smoking, antismoking socialization, and smoking onset among children. *Addictive Behavior* 22: 107-114.

[31] Khwaja, A., Silverman, D., and Sloan, F., (2007) Time Preference, Time Discounting, and Smoking Decisions. *Journal of Health Economics*, 26: 927-949.

[32] Lindbeck, A., Nyberg, S., (2006) Raising Children to Work Hard: Altruism, Work Norms, and Social Insurance. *Quarterly Journal of Economics*, 121:1473-1503.

[33] Lizzeri, A., and Siniscalchi, M., (2008) Parental Guidance and Supervised Learning. *Quarterly Journal of Economics* 123 (3): 116195.

[34] Longacre, M.L., Adachi-Meja, A.M., Titus-Ernstoff, L., Gibson, J.J., Beach, M.L., Dalton, M.A. (2009) Parental Attitudes About Cigarette Smoking and Alcohol Use in the Motion Picture Association of America Rating System. *Archives of Pediatrics and Adolescent Medicine*, 163:218-224.

[35] Morris, S., and Shin, H.S., (1998) Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks. *American Economic Review*, Vol. 88, No. 3 pp. 587-597

[36] Morris, S., and Shin, H.S., (2003) Global Games: Theory and Applications, in *Advances in Economics and Econometrics*, the Eighth World Congress (edited by M. Dewatripont, L. Hansen and S. Turnovsky), Cambridge University Press.

[37] Panebianco, F. and Verdier, T., (2015) Paternalism, cultural transmission and diffusion on complex networks. CEPR Discussion Papers.

[38] Patacchini, E. and Zenou, Y., (2016) Social networks and parental behavior in the intergenerational transmission of religion. *Quantitative Economics*, 7: 969-995.

[39] Pichler, M., (2010) The Economics of cultural formation of preferences. Mimeo.

[40] Saéz-Martí (2011) Swimming with the current? Mimeo.

[41] Saéz-Martí, M., and Zenou, Y., (2012) Cultural transmission and discrimination. *Journal of Urban Economics*, 72:137-146.

[42] Tabellini, G., (2008) The scope of cooperation: values and incentives. *Quarterly Journal of Economics*, 123:905-950.

[43] Tasker, F., and Golombok, S., (1997) *Growing up in a lesbian family*. New York: Guilford.

[44] Webley, P., and Nyhus, E., (2006) Parents influence on childrens future orientation and saving. *Journal of Economic Psychology*, 27: 140-164.

[45] Weinberg, B.A., (2001). An Incentive Model of the Effect of Parental Income on Children. *Journal of Political Economy* 109: 266-80.

[46] Wilhelm, M.O., Brown, E., Rooney, P.M., Steinberg, R., (2008) The intergenerational transmission of generosity. *Journal of Public Economics*, 92: 2146-2156.