Parental Guidance and Supervised Learning

Alessandro Lizzeri Marciano Siniscalchi*

November 30, 2006

Abstract

We propose a simple theoretical model of supervised learning that is potentially useful to interpret a number of empirical phenomena. The model captures a basic trade-off between *sheltering* the child from the consequences of his mistakes, and allowing him to *learn from experience*. We characterize the optimal parenting policy and its comparative-statics properties. We then show that key features of the optimal policy can be useful to interpret provocative findings from behavioral genetics.

1 Introduction

Humans spend a substantial fraction of their lifetimes learning how to become productive members of society. In modern advanced economies, as much as a third of the population only becomes productive after spending more than a quarter of its lifetime in some learning activity [50]. Furthermore, data from modern hunter-gatherer societies (see for instance Robson et al. [57]) suggests that protracted learning has been a common feature of human societies throughout history.

Learning has been exhaustively investigated in theoretical models by economists, statisticians, and psychologists.[1] However, these studies typically abstract from the fact that learning takes place under the supervision of parents, caregivers, teachers, advisors, and other experts for a considerable fraction of an individual's life. The economics literature has developed several models of investment in child quality, starting at least with Becker

*We thank Daron Acemoglu, Joe Altonji and Greg Duncan. Financial support from NSF Grants SES-0452317 (Lizzeri) and SES-0453088 (Siniscalchi) is gratefully acknowledged. Aaron Hauptmann provided excellent research assistance.

[1] In economics there is a vast literature that studies learning from various points of view: from Bayesian learning to adaptive learning to fictitious play. On theory, see e.g. Fudenberg and Levine [36]. Camerer [23], chapter 6 discusses both theory and experiments.

[7]). Moreover, an enormous literature in developmental psychology addresses the effects of parental care on the development of children.[2] Yet, formal modeling of *supervised learning*, i.e. the relation between parental behavior and its effects on children's learning processes, is almost absent. This is the focus of the present paper.

We develop a simple theoretical model of supervised learning that is potentially useful to address a number of empirical phenomena. The model captures a basic trade-off between *sheltering* the child from the consequences of his mistakes, and allowing him to *learn from experience*. We characterize the optimal parenting policy and its comparative-statics properties; we then argue that key features of the optimal policy may be useful to interpret provocative findings from behavioral genetics [19, 52, 56]. Before we discuss the details, we emphasize that we view this model as a first step: we acknowledge its limitations in Section 1.2. In followup work, we plan to develop complementary approaches to study additional phenomena; some of these approaches are discussed in Section 6.

1.1 Motivation

Parental care is essential for the survival of infants and young children; indeed, this is the case for all mammals and many other animal species. However, there are vast differences in 'parenting strategies' across human societies. The structure of human families and the relation of parents to their children has significantly evolved through the course of human evolution.[3] Even in present-day societies, anthropologists have documented considerable heterogeneity in the way parents interact with their children.[4] Developmental psychologists have also provided a classification of parenting styles for Western societies.[5] Finally, as

[2]It is impossible to be exhaustive in providing references. See Shonkoff and Deborah Phillips eds. [62] for a recent overview of the field with respect to early childhood development.

[3]See Lancaster and Lancaster [49] for a discussion of the very different parent/child relation that must have existed in hunter-gatherer societies relative to that in societies based on agriculture, and to that of modern parents.

[4]For example, Lamport-Commons and Miller [48], and Richman *et al.* [55] describe parental practices among the Gusii Kenyians and Highland Mayans, and highlight stark differences as compared to behavior in the typical Western family. Blurton-Jones [18] describes studies of two populations of foragers of the sub-Saharan savanna: the !Kung and the Hadza. He documents that the former appear to adopt a much more protective and close relation with their infants as compared to the latter. Interestingly, this author suggests that this difference in parenting styles may be partly related to the harsher conditions of the land inhabited by the !Kung, and that this also leads to large differences in fertility between the !Kung and the Hadza.

[5]For instance, authoritative parenting, which is characterized by high levels of warmth, support and control is thought to be more effective than either authoritarian (which lacks warmth) or permissive (which lacks control) styles of parenting. See for instance the seminal study by Baumrind [5], and also Collins et al. [28] and Demo and Cox [31] for reviews of this literature.

regards mammals and other animals, there is an enormous amount of cross-species variation in the degree to which parents invest resources in their young.[6]

The literature on the consequences of parenting is virtually unanimous in recognizing that parental support is essential for functional development in extreme situations. Harlow and coauthors[7] separated infant monkeys from their mothers; these subjects developed severe emotional and cognitive problems. The discovery of children in Romanian orphanages, who were raised with very little human contact, provided a tragic counterpart to these studies, leading to similar conclusions. These children were in the third to tenth percentile for physical growth, and "grossly delayed" in motor and mental development.[8]

However, views on the impact of differences within the 'normal' range of variation in parenting are sharply divided. A sizable literature across the social sciences argues that these differences can have important effects. Specifically, socialization researchers provide numerous studies that attempt to relate variation in parenting styles with measures of adjustment.[9]

The literature associated with *behavioral genetics* reaches a different conclusion.[10] The starting point of this literature is a criticism of much of socialization research for failing to recognize that the correlations between parenting styles and children's outcomes could be due to the shared genes between parents and their biological children. Behavioral geneticists then attempt to isolate the effects of the genes through two complementary approaches. The more direct one is to compare twins raised together by their biological parents with twins raised apart by different adoptive families.[11] One of the problems with this approach is the limited number of twins raised apart. The other approach is to compare outcomes for children with varying degrees of genetic and environmental relatedness and employ statistical techniques to estimate the percentage of variation in personality traits that is explained by genetic factors. The findings of this literature paint a very different picture of the effects of parents. For instance, behavioral genetics studies consistently find that *twins reared together are just as similar as twins reared apart.* In fact, some studies even find that twins reared together are *less similar* than twins reared apart. More generally, this literature argues that, once one controls for genetic factors, the impact of traditional measures of family environment on most personality traits is greatly

[6]See for example Clutton-Brock [27].

[7]See for instance Griffin and Harlow [39] and Harlow and Zimmerman [40].

[8]For an interesting analysis of these children, see Chisholm [26].

[9]For a survey, see for instance Collins et al. [28] and Demo and Cox [31].

[10]For an overview of this literature, see Plomin et al. [53] or Reiss et al [54]. Some of the findings of this literature have been popularized by Harris [42], [43], Pinker [51] and Ridley [56].

[11]There are several parallel projects that gather information on this front. The first, large-scale project of this kind was the Minnesota twin study: see Bouchard et al. [19].

diminished. This broad conclusion is subject to two qualifications: first, there is evidence that improvements in the family environment has positive effects on children's cognitive ability, if one restricts attention to families of low socioeconomic status (Turkheimer et al. [66]). Second, and of more immediate relevance to the present paper, recent studies suggest the intriguing possibility that parental intervention may actually respond to specific, genetically-determined traits of the child, and thus reinforce or attenuate them (Reiss et al [54]). Thus, family environment may have significant effects, even though its impact is "genetically mediated."

The behavioral-genetics literature typically focuses on personality traits and measures of cognitive achievements, and does not explore outcomes such as educational attainment or earnings. In the economics literature, the degree to which a child's home environment supports learning (as measured e.g. by how often the mother reads to her child, or whether she helps him learn numbers) has been shown in some studies to have significant effects on cognitive achievement.[12] Recent contributions by Sacerdote [60, 61], Bjorklund, Lindahl and Plug [16] and others are methodologically closer to the behavioral-genetics literature. For instance, Sacerdote [61] analyzes a sample of Korean children randomly assigned to American adoptive families; he finds that maternal education has a significant positive effect on the educational attainment of adopted children, but a much larger effect on that of biological children. Bjorklund et al. [16] analyze Swedish adoption data and report significant effects for both adoptive and biological parents. Moreover, they find evidence for a positive interaction effect between postbirth environment (nurture) and prebirth factors (nature): as we discuss in Sec. 3, this is consistent with our approach.

In our model of supervised learning, parents can *potentially* exert a significant influence on childrens' learning outcomes. However, key features of the optimal parenting policy have a number of empirical implications that can be related to the behavioral-genetics literature and data. For instance, while parenting does have an effect on learning outcomes, plausible assumptions lead to twins reared together being just as similar, or even less similar, on average, than twins reared apart. This and other findings are further discussed in Sec. 3.

The following subsection fleshes out the details of our approach, and formalizes our notion of "supervision."

[12] See e.g. Carneiro, Heckman, and Masterov [24] and Todd and Wolpin [65]. Also, a sizable literature investigates the effects of maternal employment on children's cognitive achievement. Results are mixed: some find that employment is detrimental (Baydar and Brooks-Gunn, [6]; Desai et. al. [32]; Belsky and Eggebeen [8], Bernal [12]), others that it is beneficial (Vandell and Ramanan [67]). See also the debate on the effects of family size and birth order (e.g., Black, Devereux, and Salvanes) [17].

1.2 Methodology and Summary of Results

Our model has four key features:

- There are two agents, the child and the parent. The parent is solely interested in the child's welfare, and is active for T periods. The child is active for $L > T$ periods.
- In each period, the child must perform some task. The parent has better information than the child about the correct way to perform it.
- The child *learns by doing*: at the end of each period, he receives a signal about the quality of his performance. However, *learning is costly*: the child's per-period payoff is lower the worse his performance.
- The parent's actions simultaneously modify (typically increase) the child's per-period payoff and distort (typically bias) the child's signal about his performance.

As in the standard learning models we build upon, our agents are Bayesian rational (i.e. they maximize expected utility). However, our model is set up so that the *child*'s learning problem is elementary, and its solution involves a simple adaptive rule. We are confident that our main findings can be generalized to a suitable class of non-Bayesian-rational adaptive learning rules. On the other hand, we realize that certain aspects of child development, and hence parenting, are probably best understood in a setting where computational, memory, or motor constraints (and the evolution thereof) play a central role. A model that reflects these considerations likely requires substantial departures from Bayesian learning. In any event, adopting a relatively "conservative" modeling approach has two advantages: first, we can rest assured that the decision-theoretic foundations of our basic models are sound and well-understood; second, and more importantly, insofar as the conclusions of our analysis deviate from the predictions of standard models, we can be sure that this can be attributed entirely to expert supervision.

Another limitation of our analysis is that we do not allow the parent to "describe" or "demonstrate" how to perform the task at hand. We only model one communication channel between the parent and the child, namely the former's intervention in the latter's learning process. We certainly do **not** wish to suggest that, in actual parent-child interactions, this really is the only open communication channel.[13] We do not even wish to suggest that we believe it to be the most "interesting." Rather, the present paper focuses

[13]Indeed, it is well known at least since the work of Bandura and coauthors [4, 3] that children can learn by imitating the behavior of others. Furthermore, a stream of literature in developmental psychology, starting in part with the work of Vygotsky [68], emphasizes the "social" aspects of child development. Similar considerations are true for animals (see e.g. Zentall and Galef [75].

on one specific channel that, as the preceding discussion suggests, is important in a variety of settings. Developing a model of supervised learning that incorporates more "direct" forms of teaching as well as imitation[14] is an intriguing direction for further research.

As we discussed in the previous subsection, the predictions of our model are helpful to interpret the evidence from behavioral genetics; the reader is referred to Sec. 3 for the specifics. We now briefly summarize other implications of our analysis.

One of the key features of our model is the *tradeoff between sheltering the child and allowing him to learn*. We assume that, to some extent, allowing the child to learn means that he must be allowed to make his own mistakes. However, mistakes are costly and the parent wishes to protect the child from making mistakes that are too severe. We study how this tradeoff is optimally resolved. We show that the optimal policy of the parent partially shelters the child, implying that learning may be slowed down by the presence of the parent, but that the child learns at a much smaller cost than if he were on his own. We also show that the optimal policy for the parent is linear in the child's bias at any date. This allows us to focus on one key parameter of the model: the intensity of parental intervention.

We then investigate the dynamics of parental intervention. We show that the intensity of intervention displays subtle time patterns. Specifically, for relatively low discount factors, the intensity of intervention is eventually decreasing. For high discount factors, however, the intensity of intervention may be increasing over time.

We then look at the effects of several parameters of interest that have a natural interpretation. Specifically, we study how parental intervention responds to the child's ability to learn, which can be captured by the child's initial precision in our model. Our analysis uncovers countervailing "short-run" and "long-run" effects. On one hand, a slower learner benefits less from experience in any given time period; this induces parents to provide more sheltering. On the other hand, loosely speaking, a slower learner requires more observations over his lifetime to correct his initial bias; this clearly makes sheltering less appealing. As a result, for relatively low discount factors, the short-run effect prevails, and parents intervene more intensively when children are slow learners; for high discount factors, the opposite becomes true. We provide a similar analysis of the effects of the complexity of the learning task, which corresponds to the variance of the environment in our model.

[14] See for instance Schlag [64] who provides a bounded rationality model with imitation; see also Apesteguia et al. [1] for a treatment with theory and experiments.

2 "Hand-Holding"

2.1 The basic model

The environment we consider extends the standard, single-person Gaussian learning model to incorporate a stylized form of parental intervention. Familiarity and tractability are its main advantages. We highlight the basic trade-off between *sheltering* and *learning*, and examine the key features of the optimal parenting policy. We then discuss some interesting comparative-statics results.

2.1.1 Setup

Agents and Horizon. The model features two agents, the child (he) and the parent (she). The child lives for $L > 1$ periods, whereas the parent is active (is able to supervise the child) for $T < L$ periods.

Actions and Payoffs The child must perform a task in every period. The real number M represents the correct way to perform the task *on average*; however, the correct way to perform the task at time $t = 1, \ldots, L$ is represented by i.i.d. normal random variables $X_1, \ldots, X_L$: specifically, for every t,

$$X_t \sim N\left(M, p_X^{-1}\right). \tag{1}$$

For example, M might represent the general task of "walking," or "riding a bicycle," disregarding adjustments for different surfaces, slopes, etc.; realizations of the r.v.'s X_t reflect adjustments necessary when walking or biking on some specific surface encountered at time t. See Section 2.1.3 for further discussion.

The parent's and the child's actions at time t are also real numbers, respectively denoted by $\bar{a}_t$ and $\bar{b}_t$. If, at time t, the parent chooses action $\bar{a}_t$, the child chooses action $\bar{b}_t$, and the correct way to perform the task is X_t, then the child incurs a loss of

$$(X_t + \bar{a}_t - \bar{b}_t)^2.$$

As we discuss below in greater detail, the parent knows M, the correct way to perform the task on average, and can also anticipate the child's choice $\bar{b}_t$: thus, we can think of the parent's action $\bar{a}_t$ as a *correction for the child's average mistake*. For instance, if the child is learning to walk, and has a tendency to lean forward, $\bar{a}_t$ may correspond to holding the child's hand or otherwise supporting him.

The parent discounts per-period losses at a rate $\delta \in (0, 1)$. The child's discount factor can also be assumed to be equal to δ for definiteness, but it actually plays no role in the analysis: conditional on what the child observes, the child's actions are independent of δ.

Information and Policies. The parent knows the value of M. Moreover, he observes each realization of X_t at the end of time t, after she and the child have acted. Thus, at the time she chooses her action $\bar{a}_t$, her information consists of $M, X_1, \ldots, X_{t-1}$. Consistently with this assumption, a policy for the parent (or, more concisely, a *parenting policy*) is a tuple $a = (a_1, \ldots, a_T)$ such that, for every $t = 1, \ldots, T$, $a_t : \mathbb{R} \times \mathbb{R}^{t-1} \to \mathbb{R}$; in particular, we assume throughout that a_t is a function of $M, X_1, \ldots, X_{t-1}$. The set of parenting policies is denoted by $\mathcal{A}$.

The child does not know M. To describe the signal observed by the child at the end of time period t, fix a parenting policy $a = (a_1, \ldots, a_T) \in \mathcal{A}$; Then, for every $t = 1, \ldots, T$, the child observes $X_t + a_t(M, X_1, \ldots, X_{t-1})$. We assume that the child does *not* observe X_t and $a_t(M, X_1, \ldots, X_{t-1})$ separately. The next subsection elaborates on this key assumption. For $t = T + 1, \ldots, L$, the child observes X_t. Hence, a policy for the child is a tuple $b = (b_1, \ldots, b_L)$ such that, for every $t = 1, \ldots, L$, $b_t : \mathbb{R}^{t-1} \to \mathbb{R}$; again, the interpretation is that $b_t(\cdot)$ depends upon the observations made up to time $t - 1$. The set of policies for the child is denoted by $\mathcal{B}$.

The child has a prior on M. Thus, from the child's point of view, M is a random variable: specifically,

$$M \sim \mathrm{N}\left(M_0, p_0^{-1}\right). \tag{2}$$

Similarly, as far as the child is concerned, Eq. (1) is the *conditional* distribution of X_t given M; the child also assumes that $X_1, \ldots, X_L$ are conditionally independent.

All of the above is common knowledge. In particular, we assume that the parent knows the child's prior.[15]

2.1.2 Key Analytical Assumptions

The parent correctly anticipates the child's behavior. On the other hand, we assume that *the child is not aware of the fact that the parent's choices influence his learning environment. Formally, he acts as if* $\bar{a}_t = 0$. We think of this as an interesting polar case that is helpful as an initial step; it captures the idea that the child mainly reacts to feedback from his incorrect execution of the task, but does not recognize that this feedback is affected by the parent's intervention. We consider an alternative, "textbook equilibrium" approach in Sec. 4.3; in particular, we show that, under natural assumptions on parameters, there exists a Bayesian Nash equilibrium that exhibits the same qualitative features as the model in this section. For further discussion, see Sec. 2.1.3.

[15] This is not particularly restrictive in this version of the model. If the parent does not know the value of M_0 but anticipates the form of the child's policy, he can learn M_0 in one period, provided he knows p_0. Otherwise, he can learn both M_0 and p_0 in two periods.

As a consequence of this assumption, when the child observes $X_t + a_t(M, X_1, \ldots, X_{t-1})$ at the end of time t, he acts as if that was actually the realization of X_t alone, and anticipates no further parental intervention. This implies that the child behaves as if he was facing a standard Bayesian learning problem with a quadratic penalty function:

$$\min_{(b_1,\ldots,b_L)\in\mathcal{B}} \mathrm{E}\left[\sum_{t=1}^{L} \delta^{t-1}[X_t - b_t(X_1,\ldots,X_{t-1})]^2 \Big| X_1,\ldots,X_{t-1}\right]. \tag{3}$$

In such an environment, as is well known, the optimal action at time t, given the realizations $X_1 = x_1, \ldots, X_{t-1} = x_{t-1}$, is simply the conditional expectation of X_t given all preceding realizations. Furthermore, the latter takes a particularly convenient form:

$$E[X_t | X_1 = x_1, \ldots, X_{t-1} = x_{t-1}] = \frac{p_0 M_0 + p_X \sum_{s=1}^{t-1} x_s}{p_0 + (t-1)p_X}. \tag{4}$$

However, due to the parent's intervention, at each time $s = 1, \ldots, t-1$, the child observes the quantity $x_s + \bar{a}_s$, where $a = (a_1, \ldots, a_T) \in \mathcal{A}$ is the parent's policy, x_s is the actual realization of X_s and $\bar{a}_s = a_s(M, x_1, \ldots, x_{s-1})$. Our assumption implies that the child behaves as if $X_s = x_s + \bar{a}_s$ for all $s = 1, \ldots, t-1$; hence, her optimal action at time t is

$$E[X_t | X_1 = x_1 + \bar{a}_1, \ldots, X_{t-1} = x_{t-1} + \bar{a}_{t-1}] = \frac{p_0 M_0 + p_X \sum_{s=1}^{t-1} (x_s + \bar{a}_s)}{p_0 + (t-1)p_X}. \tag{5}$$

It is useful to introduce notation that clarifies the dependence of the child's optimal policy on the parent's own policy, and simplifies the statement of our main result. Given a parenting policy $a = (a_1, \ldots, a_T) \in \mathcal{A}$, for every $t = 1, \ldots, L$, define the random variable

$$M_t^a = \begin{cases} \frac{p_0 M_0 + p_X \sum_{s=1}^{t} [X_s + a_s(M, X_1, \ldots, X_{t-1})]}{p_0 + t p_X} & t \le T \\ \frac{p_0 M_0 + p_X \sum_{s=1}^{T} [X_s + a_s(M, X_1, \ldots, X_{t-1})] + p_X \sum_{s=T+1}^{t} X_s}{p_0 + t p_X} & t > T \end{cases} \tag{6}$$

With this notation, given a parenting policy $a = (a_1, \ldots, a_T) \in \mathcal{A}$, the child's optimal policy $(b_1^a, \ldots, b_L^a) \in \mathcal{B}$ satisfies

$$b_t^a(X_1 + a_1(M), \ldots, X_{t-1} + a_{t-1}(M, X_1, \ldots, X_{t-1})) = M_{t-1}^a. \tag{7}$$

Or, more concisely, if we regard each b_t^a as a random variable, $b_t^a = M_{t-1}^a$.

Comparing Eqs. (4) and (5) immediately shows that parental intervention distorts the child's learning process. On the other hand, parental intervention directly affects the child's per-period payoff; in particular, the expected time-t penalty conditional upon the parent's information has a simple "variance plus bias" representation:

$$E[(X_t + a_t - M_{t-1}^a)^2 | X_1, \ldots, X_{t-1}, M] = p_X^{-1} + (M + a_t - M_{t-1}^a)^2. \tag{8}$$

Thus, Eqs. (5) and (8) reflect the basic tradeoff in this model.

The parent's objective is to minimize the discounted expected loss to the child. Thus, in light of the above arguments, the parent's problem is

$$\min_{a=(a_1,\ldots,a_T)\in\mathcal{A}} \mathrm{E}\left[\sum_{t=1}^{T}\delta^{t-1}(X_t+a_t-M_{t-1}^a)^2+\sum_{t=T+1}^{L}\delta^{t-1}(X_t-M_{t-1}^a)^2\Big|M,X_1,\ldots,X_{t-1}\right]. \tag{9}$$

To further clarify, this objective function corresponds to the assumptions that the parent (i) is rational, (ii.a) believes that the child is rational, and (ii.b) believes that the child expects $a_t = 0$ for all t.

2.1.3 Discussion of the Assumptions

As noted in the Introduction, the models we describe are intended to capture aspects of *supervised learning-by-doing*. The following two modelling assumptions and devices deserve additional discussion.

Randomness. In the models considered here, whenever the child attempts to perform a task, the feedback he receives (as well as the current-period outcome) is in part determined by the realization of some random process. This is consistent with standard models of learning: the feedback obtained at time t is viewed as a signal that is informative about the right way to perform a task, but does not perfectly reveal it. The learning "problem" would otherwise be trivial.[16]

The child's awareness of the parent. The effects of parental intervention upon the child's learning process may be modeled in a variety of interesting ways. Throughout most of this paper, we assume that *the child does not draw inferences about the correct way to perform a task from the parent's intervention.* It is worth emphasizing that this does **not** necessarily imply that the child is "unaware" of the presence, or even of the actions of the parent.

We regard this as an interesting polar case. In practice, the child may understand, at least in part, how parental intervention is aimed at influencing his learning process. In

[16] Randomness may be viewed as arising out of *task heterogeneity.* For example, "riding a bycicle" may be regarded as a single task, although it requires slight adjustments depending upon the type of terrain. Consequently, a child who is learning to ride on different surfaces will receive different "signals" about his ability to perform this task. Alternatively, randomness may arise out of *trembles.* In the example, a beginning rider who falls from his bike may be unable to tell whether this happened because, say, he tried to swerve too hard, or instead because he tried to swerve gently enough, but lost control of his bike and ended up swerving hard. More generally, randomness may be viewed as a "reduced-form" representation of the *complexity* of the learning task.

the opposite polar case, the child "knows" the parent's policy, and draws inferences from it, consistently with "textbook" equilibrium analysis. As noted above, we discuss results under the equilibrium, or "full awareness" assumption, in Sec. 4.3. The main findings are that (1) under natural assumptions on parameters, one Bayesian equilibrium of the game has the same qualitative features as the solution we consider here; and (2) there is a multiplicity of equilibrium outcomes.

However, the "limited awareness" assumption we adopt for most of this paper is, in our opinion, relevant and interesting in its own right. First, although children may partially understand the teaching objectives of their parents, it seems unlikely that they learn mainly by "second-guessing" them (i.e. drawing inferences from parental intervention).[17] Thus, it seems interesting to focus on effects of parental intervention on learning that are not mediated by the child's understanding of the parent's long-term strategy.

Second, while from a game-theoretic point of view, "limited awareness" is inconsistent with a "textbook" equilibrium approach,[18] it may be problematic to justify the "full awareness" assumption implicit in equilibrium analysis, in a way that reflects the essential features of actual supervised-learning scenarios.[19] On the other hand, our analysis of "limited awareness" is consistent spirit of the learning approach to games: we specify how the child learns from observations, and derive behavioral implications from our learning assumptions.[20]

The "full awareness"/"textbook equilibrium" approach will probably be useful when considering variations of our model wherein the parent's and the child's payoffs do not coincide. In the presence of conflicts of interest, *signalling* considerations are likely to be

[17] For instance, evidence exists to the effect that infants and young children actively and purposefully interact with their companions; however, these exchanges seem to be largely confined to seeking their assistance in achieving immediate goals: cf. Rogoff [58, pp. 706-708].

[18] To clarify, the "textbook" analysis of the model we consider would regard M as a payoff-relevant parameter, drawn from a normal distribution with mean M_0 and precision p_0, which is observed only by the parent. Strictly speaking, our limited-awareness assumption is inconsistent with the joint assumptions of (1) Bayesian Nash equilibrium strategies and (2) a common prior on M. However, it *is* consistent with Bayesian Nash equilibrium in a suitable Harsanyi type space featuring richer interactive hierarchical beliefs (i.e. "epistemic types": see Harsanyi [45]); in particular, it can be assumed that players' beliefs about M are consistent with a common prior, but their beliefs about each other's beliefs are not.

[19] Continuing the analysis in the previous footnote, assuming that M is drawn from some normal distribution is merely a modeling device that reduces true *incomplete* information to *imperfect* information. Conceptually, however, there is only one relevant, or "true" value of M. The appropriate learning model for situations of this kind requires assuming that parent and child engage in repeated interactions, holding the value of M fixed. But, under reasonable learning assumptions, if the child does learn the parent's strategy, he also ultimately learns how to perform the task.

[20] Furthermore, the learning and parenting policies we consider are consistent with common belief of rationality; thus, it would be inappropriate to characterize the child's belief as "unsophisticated."

essential. We propose to address these issues in future research.

No direct communication. As noted in §1.2, we do not model direct communication between the parent and the child, because we wish to focus specifically on supervised learning-by-doing.

One way to construct a model that incorporates both aspects of parent/child interaction in a textbook-equilibrium setting is to assume that the parent's objective function does *not* coincide with that of the child, and introduce a round of cheap talk before supervised learning begins. For a simple specification, continue to assume that the child's loss at time t is $(X_t + a_t - b_t)^2$, and suppose that the parent's loss is $(X_t + a_t + D - b_t)^2$. The parameter $D \in \mathbb{R}$ can be interpreted as an *intrinsic* difference between the child's and the parent's preferred way to perform the task at hand, which is distinct from the child's *informational* bias $M_0 - M$.

The analysis and results in this section can be easily adapted to this model; the same is true of our characterization of textbook-equilibrium behavior in Sec. 4.3. If we now add one round of cheap talk, the analysis of Crawford and Sobel [30] implies that full information revelation will *not* obtain in equilibrium. Thus, there will be scope for both verbal communication (cheap talk) and supervised learning.

2.1.4 Three benchmark parenting policies

Before we analyze the solution to the problem in Eq. (9), it is useful to consider three reference, or benchmark parenting policies.

Letting Go: $a_t = 0$. This is the simplest policy. Clearly, it does not induce any bias in the child's learning process.

Full Sheltering: $a_t = M^a_{t-1} - M$ (a.k.a. "The Italian Mom"). Since the parent knows the value of M, this policy minimizes the per-period loss at times $t = 1, \ldots, T$: this can be seen from Eq. (8). Intuitively, recall that the child's choice at time t is $b_t = M^a_{t-1}$, and the loss is $(X_t + a_t - b_t)$: thus, by chooosing $a_t = M^a_{t-1} - M$, the parent "shifts" the mean of X_t so that it coincides with the child's choice. In other words, the parent makes sure that the child "gets it right" on average. Of course, this has negative consequences in terms of learning: the child's belief that M^a_{t-1} is the mean of X_t is reinforced, no matter how close or distant from the true mean M it may be.

The Boot Camp: $a_t = \frac{p_0 + (t-1)p_X}{p_X}(M - M^a_{t-1})$. This policy ensures that, at the end of time t (i.e. after observing X_t), the child's posterior M^a_t will be equal to M on average.

Formally, from Eq. (29),

$$\begin{aligned}
&\mathrm{E}\left[M_t|X_1,\ldots,X_{t-1},M\right] = \mathrm{E}\left[\frac{[p_0+(t-1)p_X]M^a_{t-1}+p_X(X_t+a_t)}{p_0+tp_X}\Big|M_{t-1},M\right] = \\
&= \mathrm{E}\left[\frac{[p_0+(t-1)p_X]M^a_{t-1}+[p_0+(t-1)p_X](M-M^a_{t-1})}{p_0+tp_X}\Big|M_{t-1},M\right] \\
&= M.
\end{aligned}$$

Intuitively, we can think of this policy as exacerbating the loss to the child for an incorrect choice, thereby accelerating learning.

Thus, the present framework allows for a range of qualitatively very different parenting strategies. Moreover, the Full Sheltering and Boot Camp policies will turn out to be useful reference points to understand the main features of the optimal solution: Full Sheltering maximizes myopic payoffs, whereas the Boot Camp policy maximizes learning.

2.1.5 Characterization and key features of the Optimal Policy

The problem in Eq. (9) can be solved by dynamic programming. It is convenient to let p_t denote the quantity $p_0 + tp_X$ for $t = 0, \ldots L$.

Theorem 2.1 *The optimal parenting policy $a = (a_1, \ldots, a_T) \in \mathcal{A}$ in Eq. 9 is*

$$a_t = \gamma_t(M^a_{t-1} - M), \tag{10}$$

$$\text{where} \quad \gamma_t = \frac{1-\delta B_{t+1}\frac{p_X}{p_t}\frac{p_{t-1}}{p_t}}{1+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2}, \quad B_{T+1} = \sum_{\tau=1}^{L-T}\delta^{\tau-1}\left(\frac{p_T}{p_{T+\tau-1}}\right)^2, \quad B_t = \frac{\delta B_{t+1}}{1+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2}. \tag{11}$$

Furthermore, $\gamma_t \in (0,1)$

The proof of this and all other results are in the Appendix. For ease of reference, the coefficient γ_t is called the (time-t) *intensity of intervention*.

We emphasize two key features of the solution. First, optimal parental intervention is a *linear function of the child's bias* in each period; this is a consequence of the assumptions of Gaussian uncertainty and quadratic per-period payoff.[21] Second, and more interestingly, the solution features *partial sheltering*: the intensity of intervention γ_t lies in $(0,1)$. This

[21] However, it can be conjectured that optimal intervention is increasing in the bias, albeit non-linearly, for suitably more general specifications of per-period rewards.

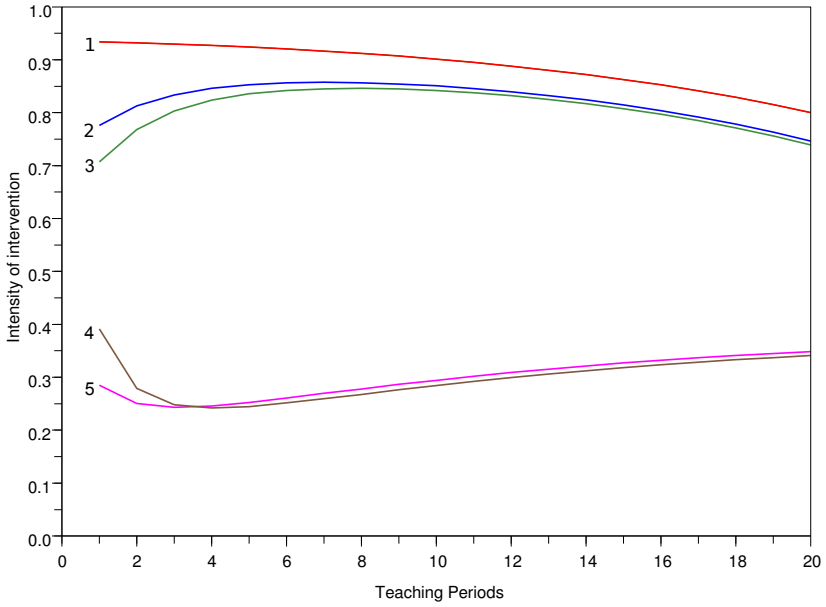

Figure 1: Intensity of intervention for different values of p_0 and p_X.

finding will play a crucial role in our analysis of the evidence from behavioral genetics in Sec. 3.[22]

It is also useful to examine a few qualitative features of the solution. Figure 1 depicts the intensity of intervention for a range of different parameters, summarized in Table 1.

Label	Color	δ	p_0	p_X
1	Red	0.9	1	0.1
2	Blue	0.9	1	1
3	Green	0.9	0.1	1
4	Brown	0.99	0.1	1
5	Magenta	0.99	1	1

Table 1: Parameters for the plots in Fig. 1. $L = 100$, $T = 20$.

[22]This second conclusion is largely dependent upon the assumption that $\delta \in (0, 1)$. This seems natural if δ is interpreted as a discount factor; however, values of δ larger than 1 might alternatively be viewed as reflecting the fact that, as the child grows and becomes an adult, his choices may have progressively more serious consequences. On the other hand, even under the current assumption that $\delta \in (0, 1)$, the length $L - T$ of the non-teaching period vis-a-vis the length T of the teaching period can *also* reflect the relative importance of mistakes in childhood and adult life.

The figures indicate that γ_t need not be a monotonic function of t. This implies that the intensity of intervention may initially rise, despite the fact that the child is becoming progressively more informed.

We can gain further intuition building on the following simple result. First, observe that the "full sheltering" and "Boot Camp" policies in the previous subsection correspond to the following values for γ_t:

$$\gamma_t^{\mathrm{FS}} = 1 \quad \text{and} \quad \gamma_t^{\mathrm{BC}} = -\frac{p_{t-1}}{p_X}. \tag{12}$$

Proposition 2.2 *In the setting of Theorem 2.1, for $t = 1, \ldots, T$,*

$$\gamma_t = \mu_t \gamma_t^{\mathrm{FS}} + (1 - \mu_t)\gamma_t^{\mathrm{BC}}, \quad \textit{where} \quad \mu_t = \frac{B_t}{\delta B_{t+1}} \in (0, 1).$$

That is, γ_t is a weighted average of a policy that maximizes learning, and one that maximizes myopic payoffs. The weight μ_t reflects the relative "cost" of biases in the current and the following periods. Specifically, the proof of Theorem 2.1 shows that the value function of the parent at time t can be written as $A_t + B_t(M_{t-1} - M)$, where B_t is the coefficient appearing in Eq. (11) and in the above Corollary. Thus, μ_t (and hence γ_t) will be higher in periods when the "cost" of mistakes is high relative to the following period.[23]

This decomposition of the intensity of intervention allows us to provide a more rigorous account of the patterns exhibited in Fig. 1. Notice that γ_t^{FS} is constant; on the other hand, since the child's posterior precision p_{t-1} increases linearly with t, the coefficient γ_t^{BC} becomes more and more negative.[24] Thus, the actual time evolution of γ_t depends upon that of μ_t. The following result contains the main tools for our analysis.

Proposition 2.3 *In the setting of Theorem 2.1, for every $T_0 \in \{1, \ldots, T\}$:*

1. *$\delta < \left(\frac{p_{T_0}}{p_{T_0+1}}\right)^2$ implies $\mu_t > \mu_{t+1}$ for all $t = T_0, \ldots, T-1$.*
2. *There exists $\delta(T_0) \in (0,1)$ such that $\delta > \delta(T_0)$ implies $\mu_t < \mu_{t+1}$ for all $t = T_0, \ldots, T-1$; furthermore, $\delta(t) \leq \delta(t+1)$ for all $t = 1, \ldots, T-1$.*

In other words, except possibly for the first few time periods, the weight μ_t (and therefore the intensity of intervention γ_t) will be *decreasing* in t for δ high and *increasing* in t for δ low. The relevant cutoffs for δ depend upon the parameters of the problem.

[23] It is immediate from Eq. (11) that $B_t < \delta B_{t+1}$.

[24] Intuitively, the child's posterior is less affected by experience in later periods, and therefore a more substantial intervention is required in order to correct a given expected bias.

To interpret, note that the impact of observations on the child's posterior is greater early on (see Eq. (6)), so sheltering in later periods induces a smaller bias. If δ is high, the parent places more weight on reducing the child's bias than on minimizing current losses, and an (eventually) increasing level of sheltering is consistent with this. We show in the next subsection that, for δ high, the parent will also want to provide little sheltering overall; thus, the optimal policy will feature low and eventually increasing degree sheltering. The brown and magenta curves in Fig. 1 exhibit these features. If, on the other hand, δ is not too high, then the opposite intuition applies: the optimal policy will feature a relatively high and eventually decreasing degrees of sheltering, as is the case for the red, blue and green curves in Fig. 1.

2.1.6 Comparative Statics

Discounting and Time Horizon. Changes in δ and L have the following consequences:

Proposition 2.4 *In the setting of Theorem 2.1, γ_t is decreasing in δ and in L.*

To interpret this result, note that, as the discount factor δ increases, and/or the number of unsupervised periods $L - T$ increases (specifically, if L increases and T is held fixed), learning the correct value of M becomes more important for the child. The Proposition confirms that, in this case, the intensity of intervention γ_t decreases.

This conclusion may appear to be at odds with the implications of typical investment models (parental or otherwise). As the discount factor increases, these models typically predict higher investment, because the balance between short-run costs and long-run benefits is tilted in favor of the latter. Here, the parent intervenes less intensely as δ increases, which would seem to suggest that parental investment is lower, not higher. But notice that less intense parental intervention translates into more costly mistakes for the child on average, and less biased learning; so, in effect, the parent allows the child to "invest more in learning," even if this entails a higher short-term cost.

Slow vs. Fast Learners The precision p_0 of the child's prior can be interpreted as reflecting the child's *speed of learning*. Refer to Eqs. (4) and (29), which characterizes Bayesian updating of the estimated mean of X_t: if p_0 is high, the child places more weight on her prior M_0 than on observations, and hence may be deemed a "slow learner" (or "stubborn").[25] If, on the other hand, p_0 is small, the opposite is true, and we may deem

[25]Note that the above does *not* imply that the child is better off if p_0 is lower: for instance, if the child happens to "be right," i.e. if $M_0 = M$, then being stubborn is beneficial; on the other hand, if M_0 is sufficiently different from M, being stubborn is harmful to the child. The interpretation of p_0 suggested in the text pertains to beliefs, not payoffs.

the child a "fast learner."

It is then interesting to ask how the child's speed of learning influences the optimal intensity of intervention. One effect is relatively straightforward: a slower learner benefits less from a reduced bias, i.e. from reduced sheltering; thus, there is an incentive to provide more sheltering, i.e. *increase* γ_t, as p_0 increases. We shall call this the "inertia" effect.

There is, however, a more subtle effect, pushing in the opposite direction. A slower learner will obviously learn slowly not just today, but also in the future; in other words, the "cost" (continuation value) of the residual bias at the end of the current period is higher for a slower learner. Thus, there is an incentive to provide less sheltering, i.e. *decrease* γ_t, as p_0 increases. We deem this the "continuation value" effect.

In order to help disentangle these effects, the decomposition in Proposition 2.2 is useful:

Proposition 2.5 *In the setting of Theorem 2.1, μ_t is increasing in p_0 for every $t = 1, \ldots, T$.*

Intuitively, this result suggests that *parents of a slower learner place more weight on sheltering than on correcting the child's bias.*

However, this does not imply that γ_t itself is increasing in p_0 for every t: indeed, the magenta and brown curves in Fig. 1 demonstrate that this need not be the case.

While Prop. 2.5 confirms that, as p_0 the intensity of intervention places more weight on the Full Sheltering intervention profile than on the Boot Camp profile, it is also the case that *the Boot Camp intervention coefficient decreases linearly in* p_0. The preceding Proposition does not indicate which of these two effects prevails.

In order to resolve this issue, we have conducted very extensive computations, exploring a large subset of the parameter space.[26] For all parameterizations of the problem, we have verified the following two properties:

1. For every $t = 1, \ldots, T$, there is $\delta_t \in [0, 1]$ such that $\frac{\partial \gamma_t}{\partial p_0} \geq 0$ iff $\delta \leq \delta_t$;
2. The cutoffs δ_t are increasing in t.

The interpretation is as follows. First, if δ is not too high, the "inertia" effect prevails, so parents of slower learners shelter more. If, however, δ is high enough, then the "continuation value" effect prevails; this is intuitive, as higher values of δ indicate that the parent cares more about the future.

Second, the monotonicity properties of the threshold imply that, as p_0 increases, either γ_t increases uniformly (as e.g. for the Green and Red curves in Fig. 1), or else the curve $t \mapsto \gamma_t$ rotates counterclockwise as p_0 increases (as is the case for the Brown and Magenta

[26] Specifically, we considered values of L from 1 to 5,000, any $T \in \{1, \ldots, L\}$, and a grid size for δ, p_X and p_0 of 0.01. Details of our computational approach, as well as Java code, are available upon request.

curves in Fig. 1), and the higher-p_0 curve crosses the lower-p_0 curve only once. Intuitively, this means that parents of a slower learner will provide *less sheltering early on, and more sheltering later*.

Task Complexity. The variance of the per-period realization X_t is a measure of the difficulty of learning the task under consideration. This is reflected both in the per-period payoff function, Eq. (8), and in the fact that observations are less informative when the variance of X_t is higher (so p_X is lower): see Eq. (2).

It is easy to see from Eq. 11 that γ_t is unchanged if p_0 is replaced with $\frac{p_0}{p_X}$ and p_X is set to 1. In other words, γ_t depends on p_0 and p_X only through the ratio $\frac{p_0}{p_X}$.

This implies that the comparative statics for p_X are exactly the reverse of the comparative statics for p_0 described above. Specifically, there are two effects: if the learning task is not very complex (p_X is high), then experience is valuable and hence the parent has an incentive to provide limited sheltering. On the other hand, if the task is easy to learn, then the child will also be able to learn it quickly later on, so providing sheltering today is not particularly costly in terms of long-term effects. As above, if δ is not too high, the first effect dominates and γ_t is decresing in p_X; and if δ is high enough, γ_t is increasing in p_X.

3 Nature and Nurture: Interpreting evidence from behavioral genetics

This section has two purposes. The more limited purpose is illustrative: we show how even this very simple and stylized model of supervised learning generates some intriguing empirical predictions. The more ambitious purpose is to use this specific model to interpret some controversial findings from behaviorial genetics. In this respect, our analysis is exploratory in nature: we view the results in this section as a starting point for future research. However, we also believe that the effects that this discussion highlights are more general and worth pointing out.

3.1 A Population Model

In order to analyze the interaction between genetic effects and parenting, we embed our simple, two-agent supervised-learning model within a population framework where parents and children are heterogeneous.

We continue to assume that the correct way to perform a task is represented by the real number M. We also continue to assume that every child has a normal prior over M. The dimension of heterogeneity we explore is the simplest one to analyze in our model:

we assume that a distribution of prior means M_0 in the population of children is given. Formally, we now treat M_0 as a random variable.

Symmetrically, we assume that parents do not observe M, and have a normal prior on it, with mean Z_0 and precision p_{Z0}. Furthermore, a distribution of the prior mean Z_0 in the population of parents is given: thus, Z_0 is also treated as a random variable.

Because M is the same for everyone in the population, the heterogeneity in priors, and in the corresponding posteriors, has payoff consequences, and can be interpreted as a a fitness measure: agents who are closer to the true mean, namely, those who are more correct on average, are better off because they make better decisions.

The exact distribution of M_0 and Z_0 is not important. We do, however, make a few assumptions relating the key uncertain quantities in the model.

First, to aid in the interpretation of this model, it is useful to think of the parent's prior beliefs $\mathrm{N}(Z_0, p_{Z0})$ as coming from (1) some prior belief that the parent held when she was born as a child, and (2) subsequent experience acquired as the parent was growing up. We also imagine that the distribution of the beliefs held by the parent as she was born as a child is the same as the distribution of the current child's beliefs. We will use one consequence of this interpetation: the variance of the parent's "prior" mean Z_0 is smaller than the variance of M_0.

We also assume that the *population* means of M_0 and Z_0 coincide with the actual, correct way to perform the task M (i.e., M_0 and Z_0 are unbiased), and that children's and parents' prior means are uncorrelated with the observations. Formally:

Assumption 1 *The joint distribution of the random variables* $(M_0, Z_0, X_1, \ldots, X_L)$ *satisfies the following restrictions.*

1. $\mathrm{Var}[Z_0] < \mathrm{Var}[M_0]$.
2. $\mathrm{E}[M_0] = \mathrm{E}[Z_0] = M$.
3. $\mathrm{Cov}[M_0, X_t] = \mathrm{Cov}[Z_0, X_t] = 0$.

Whenever we consider more than one parent and/or more than one child, Assumption 1 will apply to each child-parent pair and to the observations made by the child in that pair. Observations made by different children will be assumed to be conditionally independent.

The correlations between two individuals' prior means play a crucial role in our analysis. Specifically, *these correlations reflect the genetic relatedness of the individuals under consideration.* Thus, for instance, the prior means of unrelated parents, or of parents and adoptive children, are uncorrelated; the prior means of dizygotic twins are perfectly correlated; and so on.

We note that there may be some ambiguity as to the interpretation of the term 'parent.' We typically mean some aggregate of the two parents, which reflects the extent to which child-rearing responsibilities are shared within the family. This only matters when quantifying the genetic relatedness of a child and his biological 'parent.' Obviously, a child and his father (mother) share 50% of the genes, but it is not obvious whether the child shares more or less than 50% of his genes with the aggregate 'parent.' However, we do not need to take a stand on the precise values here; all we need is that the child have substantial (but not perfect) genetic relatedness with his biological 'parent.' Of course, a child will have zero relatedness with adoptive 'parents'; also, there is no ambiguity in the interpretation of relatedness for siblings.

An important feature of the linear-quadratic-Gaussian framework adopted here is that, even if parents do not know M, *the optimal policy has the same structure as in Theorem 2.1.* The only difference is that, at each time t, M is replaced with its conditional expected value at that time, given the parent's prior mean Z_0 and the realizations of the signals $X_1, \ldots, X_t$. We summarize these facts in the following Theorem.[27]

Theorem 3.1 *Suppose the parent does not observe M, and her prior on it is normal, with mean Z_0 and variance p_{Z0}. Then, for $t = 1, \ldots, T$, the optimal parenting policy is $a_t = \gamma_t(M_{t-1} - Z_{t-1})$, where*

$$p_{Zt} = p_Z + tp_X, \qquad Z_t = \frac{p_{Zt-1}Z_{t-1} + p_X X_t}{p_{Zt}},$$

and γ_t is given by Eq. (11) in Theorem 2.1. In particular, $\gamma_t \in (0, 1)$.

3.2 Variance Decomposition

The simplest approach to comparing the relative weight of genetic and environmental factors in determining a given personality trait is to decompose its variance into separate additive terms. We briefly discuss this approach, mainly in order to introduce the standard terminology and concepts in this literature.

The modal findings in behavioral genetics attribute approximately 50% of the variability to genetic factors (depending on the specific trait under consideration), and the remainder to environmental influences and measurement error. In particular, "shared" influences (i.e. effects that are common to all siblings in a given family) appear to be small; also, while quantitatively larger, "non-shared" environmental influences appear to have a significantly smaller effect than genetic components.

[27] Since the proof is similar to that of Theorem 2.1, we only provide a sketch in the Online Appendix.

The predictions of our model are qualitatively in line with these findings. Suppose that the parent is only active at time $t = 1$, and the child is active at times $t = 1$ and $t = 2$. From Theorem 3.1, the child's posterior at the end of time 1 is

$$M_1 = \frac{p_0 + \gamma_1 p_X}{p_1} M_0 + \left(\frac{p_X}{p_1}\right) [X_1 - \gamma_1 Z_0].$$

Its variance is easily computed under the Assumptions described above:

$$\begin{aligned} \mathrm{Var}\left[M_1\right] &= \left(\frac{p_0 + \gamma_1 p_X}{p_1}\right)^2 \mathrm{Var}\left[M_0\right] + \left(\frac{p_X}{p_1}\right)^2 \mathrm{Var}\left[X_1\right] + \\ &+ \left(\frac{p_X \gamma_1}{p_1}\right)^2 \mathrm{Var}\left[Z_0\right] - 2\frac{p_0 + \gamma_1 p_X}{p_1}\frac{p_X \gamma_1}{p_1}\mathrm{Cov}\left[M_0, Z_0\right]. \end{aligned} \tag{13}$$

Purely "genetic" effects correspond to M_0 in the model, whereas the parent's prior is a "shared environmental" effect from the point of view of the child. Finally, the signal X_1 fits the description of a "non-shared environmental" effect. We defer the dicussion of the covariance term to the next two subsections, where it plays a key role.

Qualitatively, since in the optimal parenting policy $\gamma_t \in (0, 1)$, this decomposition exhibits the main features of the evidence from behavioral genetics described above: the child's prior (genes) contributes most of the variance, followed by the signal (non-shared environment) and the parent's prior (shared environment).

3.3 A "twin experiment"

A more informative approach to evaluating genetic and environmental influences on personality traits involves studies of twins. The basic idea is to compute the correlation between personality measures for twins reared together (by their biological parents) and twins raised apart (by *different* adoptive parents). If environmental influences are the main determinants of a trait, then twins reared together should be "more similar" (i.e. exhibit higher correlation) than twins reared apart; in particular, this would be the case if *shared* environmental influences (i.e. aspects of family relationships that affect both twins equally) were the main determinant of the personality trait being examined.

However, behavioral genetics studies consistently find that *twins reared together are just as similar as twins reared apart* (Plomin and Daniels, [52]; Bouchard et al. [19]). In fact, some studies even find that twins reared together are *less similar* than twins reared apart. This latter finding seems paradoxical: after all, if shared influences had *no* effect, one would expect to find no difference at all between twins raised together or apart.

Our model can generate these patterns, but suggests an interpretation that differs from those typically offered in the behavioral-genetics literature. We continue to consider a

two-period environment. Consider two twins, a and b, with prior and posterior means M_{0i} and M_{1i} respectively. We examine first the case of twins raised *apart* by adoptive parents. Denote the prior of child i's parent by Z_{0i}. We assume that Z_{0a} and Z_{0b} are independent and identically distributed;[28] their common variance is denoted by $\mathrm{Var}[Z_0]$. Then, from Theorem 3.1,

$$M_{1i} = \frac{p_0 + \gamma_1 p_X}{p_1} M_{0i} + \left(\frac{p_X}{p_1}\right)[X_{1i} - \gamma_1 Z_{0i}], \tag{14}$$

and we can compute the covariance between the twins' posteriors as follows:

$$\mathrm{Cov}[M_{1a}, M_{1b}] = \left(\frac{p_0 + \gamma_1 p_X}{p_1}\right)^2 \mathrm{Cov}[M_{0a}, M_{0b}]. \tag{15}$$

Notice that there is no term involving covariances between parents and children: the assumption that the twins are raised by different adoptive families implies that these covariances are all zero.[29]

Next, consider twins raised by their biological parents; denote the parent's prior mean by Z_0. Then each twin's posterior can be computed as in Eq. (14) above, except that Z_{0i} is replaced with Z_0. The covariance between the twins' posteriors is given by

$$\begin{aligned} \mathrm{Cov}[M_{1a}, M_{1b}] &= \left(\frac{p_0 + \gamma_1 p_X}{p_1}\right)^2 \mathrm{Cov}[M_{0a}, M_{0b}] + \left(\frac{p_X \gamma_1}{p_1}\right)^2 \mathrm{Var}[Z_0] \\ &- 2\frac{p_0 + \gamma_1 p_X}{p_1}\frac{p_X \gamma_1}{p_1}\mathrm{Cov}[M_{0a}, Z_0] \end{aligned} \tag{16}$$

(we use the fact that $\mathrm{Cov}[M_{0a}, Z_0] = \mathrm{Cov}[M_{0b}, Z_0]$, because twins share 100% of their genes). It follows that the difference between the covariance for twins raised together and for twins raised apart is

$$\left(\frac{p_X \gamma_1}{p_1}\right)^2 \mathrm{Var}[Z_0] - 2\frac{p_0 + \gamma_1 p_X}{p_1}\frac{p_X \gamma_1}{p_1}\mathrm{Cov}[M_{0a}, Z_0]. \tag{17}$$

Genetic considerations imply that a child's prior is positively correlated with her parent's prior. Also, recall that, by Assumption 1, $\mathrm{Var}[M_{0a}] > \mathrm{Var}[Z_0]$, which implies that $\mathrm{Cov}[M_{0a}, Z_0] = \mathrm{Corr}[M_{0a}, Z_0]\sqrt{\mathrm{Var}[M_{0a}]\,\mathrm{Var}[Z_0]}$. Finally, recall that the optimal parenting policy prescribes that $\gamma_t \in (0, 1)$. Taken together, these observations imply that the difference between the covariances of the twins' posteriors in the two scenarios is likely to

[28] Bouchard et al. [19] argue that this is the empirically relevant case.

[29] As argued by Bouchard et al. [19] and Reiss et al [54], the evidence is indeed that the twins are randomly assigned in the population, so that these covariances can be reasonably thought to be negligible in the data.

be small (for instance, relative to the covariances themselves). In fact, if the child's prior mean M_{0a} is sufficiently dispersed in the population, then our model predicts that twins raised together will be less similar than twins raised apart. Appendix A.2.4 provides simple conditions under which this is the case, under the assumption that the parent's mean Z_0 is also the result of Bayesian updating.[30]

It should also be emphasized that our conclusions rely on the fact that $\gamma_t \in (0,1)$ in the optimal parenting policy. If the parent adopted the "Boot Camp" policy $\gamma_1^{BC} = -\frac{p_0}{p_X}$ (or a similar policy), then the above difference would be strictly positive. If the parent adopted the "letting go" policy $\gamma_t = 0$, then the difference would be exactly zero.

The intuition for the mechanism that generates this conclusion in our model is simple. Each twin raised by adoptive parents is likely to be less similar to her parent than if she was raised by her biological parent. On average, this will lead adoptive parents to provide *more sheltering*. This is consistent with evidence from a variety of sources. For instance, Hoopes [44] finds that "adoptive mothers are more protective and careful with the children... adoptive mothers and fathers reported that they fostered more dependency than the biological fathers and mothers. The latter group admitted to greater feelings of irritability regarding their children, and the fathers tended to force independence, suppress affection, and accelerate development (p.23)." Furthermore, these more protective attitudes of adoptive parents "may have their effect on the children, who, at 5 years of age, were rated as a little less confident and less willing and attentive in task completion (p.27)." Warren [71] shows that "adoption significantly increases the likelihood of referral for psychiatric treatment, even after controlling for the fact that adoptees are significantly more likely to be referred when they display few problems."

But this means that non-shared environmental influences will have fewer opportunities to affect the twins' posterior, which, as a result, will be *more similar to their prior, and hence more similar to one another* on average. Conversely, twins raised by their biological parents will be subject to less sheltering, because their priors are positively correlated with their parents'. Hence, their non-shared experiences will have a greater role, and they will end up being less similar to one another.

We believe that this key effect does not rely on the specifics of our model. For instance, we could obtain similar effects if we assumed that intervention by parents is solely driven

[30] This "extreme" prediction depends, among other factors, on the assumption (maintained thoughout this section) that the observations X_{1i} made by twins reared together are just as correlated as those made by twins reared apart; in particular, they are conditionally independent. If instead twins reared together were exposed to more similar (i.e. correlated) observations, this would tend to offset the effect described above; the resulting predicted covariances would thus be more in line with the empirical findings of these studies, even if the population means are widely dispersed.

by a desire to have their children be similar to them.[31] Under this assumption, because the biological parents are more similar to their children than the adoptive parents, the latter intervene more intensely, thus generating similar effects to those discussed above.

The observation that adoptive parents provide more sheltering has a further empirical implication in our model: the ability of adopted children to perform the task under consideration is more dispersed—that is, the variance of the posterior mean is larger for adopted children. This is because adopted children are less exposed to non-shared experiences, and hence don't learn the correct way to perform the task at hand as fast as children reared by their biological parents.

This can be seen from Eq. (14); also refer to the variance decomposition in Eq. (13), which applies here, too, with the obvious notational changes. For a child reared by adoptive parents, $\mathrm{Cov}\left[M_{0i}, Z_{0i}\right] = 0$, due to lack of genetic relatedness;[32] on the other hand, $\mathrm{Cov}\left[M_{0i}, Z_{0i}\right] > 0$ for a child reared by her biological parents. Finally, this covariance is multiplied by a negative constant because $\gamma_1 \in (0, 1)$; this delivers the required conclusion.

However, it is important to point out that this is a statement about the *variance* of posterior means, not their population average. In fact, Eq. (14) implies that the average posterior mean will be the same for adopted and biological children. This is a consequence of the linearity of the parent's optimal policy.

3.4 Genetic Cascade

Another way to investigate the interplay of genetic factors and parenting is to compare the correlations between personality traits of pairs of children that differ in the fraction of genes they share. For example, one may compare the correlations for monozygotic twins (who share 100% of the genes), dizygotic twins and full siblings (who share 50% of the genes), half siblings (who share one parent, and hence only 25% of the genes), and biologically unrelated children. The collection of these correlations, ordered in terms of genetic relatedness, is called a *genetic cascade* (Reiss et al. [54], p. 48).

In our model, we can compute the correlation between the posterior means of two children raised by the same parents. As noted above, different degrees of genetic relatedness correspond to different correlations between the children's prior means. The mapping from prior to posterior correlations plays the role of a "genetic cascade."

For $i = a, b$, M_{0i} denotes child i's prior, and M_{1i} is child i's posterior; similarly, X_{1i}

[31] Bisin and Verdier [13, 14] make such an assumption in a model of cultural transmission.

[32] And, if Z_{0i} is interpreted as a Bayesian update as well, also due to the assumed independence of an individual's prior mean with the observations X_t.

denotes i's observation (non-shared experience) at time 1. Then, from Theorem 3.1,

$$M_{1i} = \frac{p_0 + \gamma_1 p_X}{p_1} M_{0i} + \left(\frac{p_X}{p_1}\right)[X_{1i} - \gamma_1 Z_0].$$

Therefore, the correlation between the two children's posteriors is

$$\begin{aligned} \mathrm{Cov}\left[M_{1a}, M_{1b}\right] &= \left(\frac{p_0 + \gamma_1 p_X}{p_1}\right)^2 \mathrm{Cov}\left[M_{0a}, M_{0b}\right] + \left(\frac{p_X \gamma_1}{p_1}\right)^2 \mathrm{Var}\left[Z_0\right] \\ &- \frac{p_0 + \gamma_1 p_X}{p_1} \frac{p_X \gamma_1}{p_1} \left(\mathrm{Cov}\left[M_{0a}, Z_0\right] + \mathrm{Cov}\left[M_{0b}, Z_0\right]\right). \end{aligned} \tag{18}$$

The resulting genetic cascade typically conforms to what one obtains in studies of relatively highly heritable traits. First, the difference between the posterior covariance for monozygotic twins and for dizygotic twins or full siblings is proportional to the difference between the respective prior covariances (which reflect their genetic relatedness): the terms involving Z_0 are common and hence cancel out. Since monozygotic twins share 100% of the genes, whereas dizygotic twins and full sibs only share 50% of the genes, it follows that monozygotic twins have more similar posteriors than do dizygotic twins and full sibs.

The difference between the posterior covariances for full sibs (or dizygotic twins) and, say, genetically unrelated siblings is typically also positive; however, the analysis is more subtle. To clarify, consider a parent with three children $i = a, b, c$; a and b are biological children (hence full siblings), and c is adoptive. Then

$$\begin{aligned} \mathrm{Cov}\left[M_{1a}, M_{1b}\right] - \mathrm{Cov}\left[M_{1a}, M_{1c}\right] &= \left(\frac{p_0 + \gamma_1 p_X}{p_1}\right)^2 \mathrm{Cov}\left[M_{0a}, M_{0b}\right] \\ &- \frac{p_0 + \gamma_1 p_X}{p_1} \frac{p_X \gamma_1}{p_1} \mathrm{Cov}\left[M_{0b}, Z_0\right]. \end{aligned} \tag{19}$$

Thus, as long as the variance of the parent's prior Z_0 is small relative to the variance of M_{0i} (as would be expected if Z_0 actually reflects the parent's past experiences), the above difference is positive.

It is worth emphasizing that the second term in the right-hand side of Eq. (19) reflects a "non-shared" aspect of parental intervention—the fact that the parent responds differently to child b's and child c's prior. However, notice that this means that the parent is responding to *genetic* features of her children. This is consistent with recent, intriguing findings in behavioral genetics: *even non-shared environmental influences are actually genetically mediated*. For instance, this is the central finding reported by Reiss et al. [54] in their influential study of parental influences on adolescent development.

On the other hand, the cascade reflects environmental effects as well, as should be expected in a model of parenting. In particular, the posteriors of blended siblings are

correlated—a measure of the effects of *shared* environmental influences. However, it should be noted that, *since* $\gamma_t \in (0,1)$ *in the optimal policy*, this covariance is likely to be quite small. Again, this is consistent with most findings in behavioral genetics, which report only modest effects of the shared environment.

4 Extensions

4.1 Costly Parental Intervention

So far we have assumed that the parent's payoff per period coincides with that of the child for any given realization of uncertainty X_t and any choice of actions $\bar{a}_t$ and $\bar{b}_t$. In particular, this implies that the only "cost" of parental intervention is via the induced distortion of the child's beliefs.

We now extend our basic model and introduce an explicit cost of parental intervention. For analytical tractability, we assume that this additional cost is quadratic in the parent's action $\bar{a}_t$. Specifically, the *parent's* loss at time t, given X_t, $\bar{a}_t$ and $\bar{b}_t$, is now

$$(X_t + \bar{a}_t - \bar{b}_t)^2 + C\bar{a}_t^2 \tag{20}$$

where $C \geq 0$. Our main result, Theorem 2.1, admits a straightforward generalization to this environment.

Theorem 4.1 *Under costly parenting, the optimal parenting policy* $a = (a_1, \ldots, a_T) \in \mathcal{A}$ *is* $a_t = \gamma_t(M_{t-1}^a - M)$*, where*

$$\gamma_t = \frac{1 - \delta B_{t+1} \frac{p_X}{p_t} \frac{p_{t-1}}{p_t}}{1 + C + \delta B_{t+1} \left(\frac{p_X}{p_t}\right)^2}, \quad B_{T+1} = \sum_{\tau=1}^{L-T} \delta^{\tau-1} \left(\frac{p_T}{p_{T+\tau-1}}\right)^2, \quad B_t = \frac{\delta B_{t+1}}{1 + \delta B_{t+1} \left(\frac{p_X}{p_t}\right)^2}. \tag{21}$$

Furthermore, $\gamma_t \in (0,1)$.

Notice that the optimal parenting policy has exactly the same features as in the costless case; the only difference is the denominator of the intensity of intervention γ_t, which now includes the cost parameter C. Thus, it is still the case that partial sheltering occurs at the optimum, so our interpretation of the evidence from behavioral genetics (Sec. 3 is preserved under costly parenting.

4.2 Simultaneous learning in the Gaussian framework

Another interesting extension is *simultaneous learning*. Children typically engage in multiple simultaneous learning processes; furthermore, experimental evidence points to *com-*

plementarities between different aspects of cognitive development.[33] It is then natural to ask how a supervisor might take advantage of these complementarities to resolve the sheltering/learning trade-off, and more generally aid in the development process.

The Gaussian framework can be readily extended to handle learning of multiple tasks. As we shall presently demonstrate, the main technical finding is that the parent's intervention in any given task is, in general, a linear function of the child's bias in *all* tasks. Preliminary numerical exploration suggests than the parent may take advantage of this in interesting and somewhat unexpected ways; this invites further investigation.

We continue to assume that the child lives for $L > 1$ periods, and that teaching terminates at time $T \in \{1, \ldots, L-1\}$ (i.e. there is at least one teaching period).

The child must now learn how to perform n tasks. Extending the notation of Section 2.1, let $M \in \mathbb{R}^n$ denote the vector describing the correct way to perform each task on average and; for every $t = 1, \ldots, T$, let X_t denote the n-dimensional random vector that describes the correct way to perform each task at time t.

The parent knows the value of M; the child has a multinormal prior over its possible values, with precision matrix p_0 and mean vector M_0. The parent and child both believe that, conditional upon M, each X_t is a multinormal vector with mean M and precision matrix p_X. Thus,

$$X_t | M \sim N\left(M, p_X^{-1}\right) \quad \text{and} \quad M \sim N\left(M_0, p_0^{-1}\right). \tag{22}$$

The parent's and child's actions at time t, denoted by $\bar{a}_t$ and $\bar{b}_t$ respectively, are also n-dimensional real vectors. The child's loss at time t is given by

$$(X_t + \bar{a}_t - \bar{b}_t)'(X_t + \bar{a}_t - \bar{b}_t).$$

We can also incorporate a quadratic cost-of-parenting term, extending the model in Sec. 4.1. Thus, the parent's loss at time t is

$$(X_t + \bar{a}_t - \bar{b}_t)'(X_t + \bar{a}_t - \bar{b}_t) + \bar{a}_t' C \bar{a}_t,$$

where C is an $n \times n$ symmetric matrix.

By analogy with the single-task case, a parenting policy is a tuple $(a_1, \ldots, a_T)$ such that $a_t : \mathbb{R}^n \times \mathbb{R}^{n \cdot (t-1)} \to \mathbb{R}^n$ for each t. That is, the parent's intervention at time t depends upon the value of the vector M and the realizations of the random vectors $X_1, \ldots, X_{t-1}$. Similarly, a policy for the child is a tuple $(b_1, \ldots, b_L)$ such that, for every t, $b_t : \mathbb{R}^{n \cdot (t-1)} \to$

[33] For instance, Gopnick and Meltzoff [38] report that children who use more names are more likely to exhaustively sort objects into different categories (see also Waxman [69]).

$\mathbb{R}^n$. We continue to denote the set of policies for the parent and the child by $\mathcal{A}$ and $\mathcal{B}$ respectively.

Finally, in accordance with the assumptions discussed in Sec. 2.1.3, given a parenting policy $(a_1, \ldots, a_T) \in \mathcal{A}$, the child's beliefs about M at the end of each time period $t = 0, 1, \ldots, L$ are characterized by the precision matrix $p_t \equiv p_0 + tp_X$ and the mean vector

$$M_t^a = \begin{cases} p_t^{-1}\left[p_0 M_0 + p_X \sum_{s=1}^{t}[X_s + a_s(M, X_1, \ldots, X_{s-1})]\right] & t \leq T \\ p_t^{-1}\left[p_0 M_0 + p_X \sum_{s=1}^{T}[X_s + a_s(M, X_1, \ldots, X_{s-1})] + p_X \sum_{s=T+1}^{t} X_s\right] & t > T \end{cases}$$

Also, the optimal policy $b^a = (b_1^a, \ldots, b_L^a) \in \mathcal{B}$ for the child, given the parent's policy $a = (a_1, \ldots, a_T) \in \mathcal{A}$, satisfies $b_t^a = M_{t-1}^a$ for all $t = 1, \ldots, L$. Again, these are immediate extensions of the corresponding results for the single-task case.

The parent still minimizes her discounted expected loss. Theorem 2.1 can then be generalized as follows.

Theorem 4.2 *For $t = 1, \ldots, T$, the optimal parenting policy $a = (a_1, \ldots, a_T) \in \mathcal{A}$ satisfies $a_t = \Gamma_t \cdot (M_{t-1}^a - M)$, where Γ_t is inductively defined as follows:*

$$\begin{aligned} B_{T+1} &= \sum_{\tau=1}^{L-T} \delta^{\tau-1} p_t p_{t+\tau-1}^{-1} p_{t+\tau-1}^{-1} p_t, \\ \Gamma_t &= [I + C + \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_X]^{-1}[I - \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_{t-1}], \\ B_t &= (I - \Gamma_t)'(I - \Gamma_t) + \delta(p_{t-1} + p_X\Gamma_t)' p_t^{-1} B_{t+1} p_t^{-1}(p_{t-1} + p_X\Gamma_t). \end{aligned}$$

4.3 "Textbook" Equilibrium analysis

With regards to equilibrium analysis of the "Hand-Holding" models, three main findings can be mentioned; we focus on the single-task environment for simplicity.

First, if parental intervention is intrinsically costless (as in our main model), then for any choice of the parameters, the following profile of strategies is a (perfect Bayesian) equilibrium: the parent chooses $a_t = M_0 - M$ for $t = 1, \ldots, T$; the child guesses M_0 in the first T periods, disregarding the realization of $X_t + a_t$, and then proceeds as in the unsupervised case from time $T + 1$ onwards. In this equilibrium the parent opts for a *fully sheltering* strategy, which however prevents learning in the first T periods: the child is aware that he is essentially observing draws from a normal distribution with mean $M + a_t = M_0$ in each period, regardless of the realized value of M, and hence he simply discards these observations. Thus, in this equilibrium, the child does not learn anything while being supervised by the parent. However, it is easy to see that this particular equilibrium disappears as soon as costs are introduced.

Second, under natural assumptions on parameters, there exists a partially sheltering equilibrium in which the parent's strategy is linear in the child's bias, as in Theorem 2.1. In particular, it is characterized by a system of equations similar to Eq. (11), except that the quantities p_t and B_t are replaced by functions of all the equilibrium values of $\gamma_1 \dots \gamma_T$.

Third, simple two-period examples demonstrate that the child's welfare can be *lower or higher* in equilibrium than under the assumptions about beliefs stated above. Intuitively, the child can make better use of the information he receives (because he can partially filter out the parent's intervention); on the other hand, this induces the parent to shelter more than she would under the original assumptions about beliefs.

To illustrate these points, we solve a two-period, equilibrium version of the basic model in Sec. 2.1 (this can be extended to arbitrarily many periods). Since a policy for the parent consists of a single function of the true mean M, we drop time indices and denote it simply by $a : \mathbb{R} \to \mathbb{R}$. Equilibrium quantities will be denoted by the subscript "e".

Begin by conjecturing a *linear equilibrium*: that is, a Bayesian Nash equilibrium where the parent's equilibrium policy is linear in the child's initial bias: $a^e(M) = \gamma^e(M_0 - M)$, where $\gamma^e \in \mathbb{R}$ is the equilibrium intensity of intervention (again dropping the time index). Then, in equilibrium, the child knows that her time-1 observation is a realization of $X_1 + a(M) = X_1 + \gamma^e(M_0 - M)$, and not just X_1; thus, she updates accordingly. The following Lemma provides the details.

Lemma 4.3 *For all $\gamma \in \mathbb{R}$,*

$$\mathrm{E}\left[M | X_1 + \gamma(M_0 - M) = x\right] = \frac{p_0 - p_X\gamma(1-\gamma)}{p_0 + p_X(1-\gamma)^2} M_0 + \frac{p_X(1-\gamma)}{p_0 + p_X(1-\gamma)^2} x.$$

In a linear equilibrium with intensity of intervention γ^e, the child's policy $b^e = (b_1^e, b_2^e)$ must then satisfy

$$b_1^e = M_0,^{34} \quad b_2^e(x) = \mathrm{E}[M | X_1 + \gamma^e(M_0 - M) = x] \; \forall x \in \mathbb{R}.$$

Consequently, a necessary and sufficient condition for the existence of such an equilibrium is that, for every $M \in \mathbb{R}$,

$$\begin{aligned} \gamma^e(M_0 - M) \quad \in \quad & \arg\max_{\bar{a} \in \mathbb{R}} \mathrm{E}\left[(X_1 + \bar{a} - M_0)^2 + \right. \\ + \quad & \left. \delta \left(X_2 - \frac{p_0 - p_X\gamma^e(1-\gamma^e)}{p_0 + p_X(1-\gamma^e)^2} M_0 - \frac{p_X(1-\gamma^e)}{p_0 + p_X(1-\gamma^e)^2} (X_1 + \bar{a}) \right)^2 \middle| M \right] \end{aligned} \tag{23}$$

[34] In any equilibrium, the child's guess at time 1 only influences her loss at time 1, because the parent must choose an action $a(M)$ without observing b_1^e. Thus, in particular, in a linear equilibrium the child's optimal choice is $b_1^e = \mathrm{E}\left[X_1 + \gamma^e(M_0 - M)\right] = M_0$.

The following Proposition and Corollary characterize the linear equilibria of this game.

Proposition 4.4 *For all parameter values, there is a linear equilibrium with* $\gamma^e = 1$. *Furthermore, if* $\delta p_X > p_0$, *there is a linear equilibrium with*

$$\gamma^e = \hat{\gamma} \equiv 1 - \sqrt{\frac{\sqrt{\delta p_0 p_X} - p_0}{p_X}} \in (0, 1).$$

Corollary 4.5 *If there is a linear equilibrium with* $\gamma^e = \hat{\gamma}$, *then there is also a payoff-equivalent linear equilibrium with* $\gamma^e = 2 - \hat{\gamma}$; *there are no other linear equilibria. Furthermore, if* $\gamma^e \in \{\hat{\gamma}, 2 - \hat{\gamma}, 1\}$ *is the intensity of intervention in a linear equilibrium, then for every* $k \in \mathbb{R}$ *there is a payoff-equivalent equilibrium with* $a(M) = k + \gamma^e(M_0 - M)$.

Thus, as noted above, under the condition $\delta p_X > p_0$, one equilibrium exhibits the essential features of the solution we consider in Sec. 2: the parent's policy is linear in the child's bias, and in particular the intensity coefficient lies between zero and one. This coefficient is also decreasing in δ, as in Proposition 2.4. Such an equilibrium exists, provided the child is not too slow to learn relative to the complexity of the task: otherwise, the parent fully shelters ($\gamma^e = 1$) in any linear equilibrium.

Also, adding a constant k to the parent's action a leads to a formally distinct but payoff-equivalent equilibrium: the child merely adjusts her guesses b_0 and b_1 so as to offset the constant k. Finally, it turns out that, if there is an equilibrium with $\gamma^e \in (0, 1)$ as in Prop. 4.4, then there is another linear equilibrium with intensity coefficient equal to $2 - \gamma^e$: intuitively, this is the "mirror image" of the original linear equilibrium; the child can also adjust her learning rule accordingly, and again one obtains a payoff-equivalent equilibrium. The above Corollary summarizes these facts.

Table 2 below shows that, depending on parameter values, the child's welfare, as perceived by the informed parent, may be either higher or lower under textbook equilibrium behavior. The figures in Table 2 suggest that, if the child's initial bias is large, then for relatively low values of p_X the child's expected loss is higher in a linear equilibrium than under the assumption that the child is unaware of the parent's intervention, whereas the opposite is true for relatively high values of p_X. However, this ranking is reversed if the child's bias is small. Additional numerical experimentation seems to confirm this pattern. Finally, the expected loss in a linear equilibrium with $\gamma^e \in (0, 1)$ may be either higher or lower than in the full-sheltering equilibrium with $\gamma^e = 1$; as one might expect, the full-sheltering equilibrium is superior for if the child's initial bias is small.

δ	p_0	p_X	$M_0 - M$	γ	Loss	γ^e	Eqm Loss	Eqm loss ($\gamma^e = 1$)
0.9	1	2	2	0.5714	3.7214	0.5867	3.8041	4.55
0.9	1	5	2	0.5385	2.7204	0.5264	2.3013	3.98
0.9	1	2	0.5	0.5714	1.3107	0.5867	1.2885	1.175
0.9	1	5	0.5	0.5385	0.6435	0.5264	0.7103	0.605

Table 2: Textbook Equilibrium and Child Welfare

5 Discussion

5.1 Additional Related Literature

Influencing other agents' beliefs is a recurring theme in information economics and game theory. The main contrast between our model and most contributions on communication in models of asymmetric information is the following. In these models, the presence of conflicting incentives either makes it impossible to credibly communicate all the information or introduces the necessity of distorting actions in order to make communication credible. In our model on the other hand, although incentives are aligned, information cannot be directly communicated but must be transmitted through a (possibly costly) manipulation of the learning process. The informed agent in our models faces very different incentives from those in standard models of asymmetric information.

A "teaching" metaphor is sometimes employed to describe strategies in the literature on learning in games: cf. e.g. Kalai and Lehrer [46], Section 2.4. However, this seems mainly an expository device to clarify features of certain strategies; players do not have any private information to convey (besides their future dispositions to play). In the absence of the "teacher," the "learner" would have nothing to learn, which is definitely not the case in supervised learning models.

The literature on social learning is also related. In particular, Smith and Sorensen [63] analyze optimal experimentation and its interaction with informational herding. Although teaching is not the focus of their paper, these authors provide a two-period example featuring a "professor" who takes the more informative of two possible actions so as to maximize the "student's" expected payoff. This intriguing example may point in the direction of a full-blown model of *imitative* learning, or "teaching by doing" on the parent's part; by comparison, our concern in this paper is "learning by doing" on the child's part, under the supervision of the parent.

Benabou and Tirole [10] consider the interplay between the provision of incentives and the self-confidence of an agent: provision of explicit incentives by an informed principal

(e.g. a parent) may be a negative signal of the ability of an agent (the child). They show that in some circumstances, explicit incentives can be counterproductive. Other authors (e.g., Rubinstein [59]) and Brunnemeier and Parker [22]) have investigated how beliefs may be optimally chosen by a 'principal' to maximize the payoffs of an 'agent'. In these settings, optimistic beliefs may be advantageous.

Ettinger and Jehiel [34] propose a model of deception in which "manipulative" agents influence the beliefs of less sophisticated, "deceived" agents. In particular, less sophisticated agents make inferences on the basis of coarser information about their opponent than is available in the game—an assumption consistent with the so-called "fundamental attribution error." This bounded-rationality assumption leads to the possibility of deception. It should be noted, however, that no deception is possible if both agents are fully rational; by way of contrast, as we show in Sec. 4.3, the main qualitative features of our analysis are consistent with textbook-equilibrium analysis.

Finally, the transmission and adoption of cultural traits had been developed by, among others, Cavalli Sforza-Feldman [25], and Boyd-Richerson [20]. In most of these literature, the transmission is independent of any of the choices of the parents. Bisin and Verdier [13], [14], and Bisin, Topa, and Verdier [15] develop models in which cultural transmission occurs as the result of the socialization effort endogenously chosen by the parent. However, the effectiveness of socialization is modeled as a reduced-form function of some effort measure.

5.2 Applications

The approach outlined in this paper may be useful to address broader issues pertaining to learning in a multi-person environment. First, at an abstract level, **schooling** may be viewed as supervising more than one simultaneous learning process, with only limited opportunities to target interventions to individual learners. For instance, in the setting of §2.1, assume that there n 'students' with different priors about M, and one 'instructor'. At every time t, each student must guess the realization of the r.v. X_t; the instructor must choose the same level of intervention a_t for all students. If $b_{t,i}$ is student i's guess at time t, his payoff is $-(X_t + a_t - b_{t,i})^2$. If students do not observe each other's actions, the model is an n-fold repetition of the model in Section 2.1, except that the instructor's choice of a_t must be the same in every "copy" of the problem. We can then analyze the impact of different objective functions for the instructor on (i) the distribution of ability in the population of students, and (ii) the ability of a single student, given the distribution of other students' abilities.

A more interesting model of **peer effects** may plausibly be obtained by allowing students to observe each other's guesses. This would likely be beneficial for "low-ability"

students; however, it might be the case that high-ability students are better off than they would be if actions are unobservable. Intuitively, in a very heterogeneous class, the instructor's choice of a_t is likely to be significantly different from the choice she would make if she was interacting with a single high-ability student; but, as low-ability students improve due to the presence of high-ability peers, the

class becomes less heterogeneous, and as a consequence even high-ability students may benefit.

6 Concluding Remarks

We have presented a model of supervised learning that focuses on the tradeoff between sheltering a child and letting him learn by allowing him to make mistakes. Clearly, the model is not suitable to deal with all aspects of parent-child interactions, or even of supervised learning. By changing the nature of the child's learning process, different effects can be explored.

For instance, the model we analyze in this paper features a rather fine-grained form of parental intervention. As an alternative, suppose that, in each time period, the parent can either allow the child to perform a task, or intervene and perform the task on the child's behalf. If the parent intervenes, she can minimize per-period loss due to her superior information; however, the child does not observe the realization of the uncertainty in the current period.[35] Thus, the parent acts as a *gatekeeper*, but cannot directly influence the child's posterior, even on average. While this model presents a number of technical hurdles to overcome, preliminary analysis suggests that it may capture different aspects of supervised learning.

Also, we have restricted attention to a particularly simple learning problem for the child. The data that the child observes is independent of his actions (except, indirectly via the parent's choice) and therefore, so is his learning. This has kept our analysis relatively tractable because the child's choice problem is essentially a static one. Another important dimension learning and therefore another potential role for parental supervision emerges when the child's learning task involves *experimentation*. It is well known that in these circumstances, an agent may suffer incomplete learning (see Easley and Kiefer [33] and Brezzi and Lai [21]), and has positive probability of choosing the incorrect action. In preliminary work (Lizzeri and Siniscalchi, in progress) we provide several alternative models in which the child faces a multi armed bandit choice problem. A number of novel phenomena emerge in this environment. For instance, we show that, if the parent supervises the child

[35] In keeping with the assumptions discussed in section 2.1.3, the child does not make inferences from the fact that the parent intervenes: he simply "misses one observation."

sufficiently long, the probability that the child chooses the wrong action goes to zero. Interestingly, however, for this to hold, the child must come to hold posterior beliefs that are incorrect. In particular, parental intervention tends to make the child *overconfident* about the right action.[36] This comes about because the parent optimally censors the learning process of the child in a way that ends up exaggerating how bad the wrong actions are and how good the right action is. In another experimentation problem, we show that the parent may have the incentive to give the child the false impression that an environment is safe so as to induce the child to engage in experimentation.

A Appendix

A.1 Proof of Theorems 2.1, 4.1 and 4.2

Note: except for the statement that $\gamma_t \in (0,1)$, Theorems 2.1 and 4.1 are special cases of Theorem 4.2. So, we shall prove the latter first, and then show that $\gamma_t \in (0,1)$.

For ease of reference, we reproduce the definition of the key vector quantity M_t^a, representing the mean of the child's posterior at the end of time t, given the parent's policy $a \in \mathcal{A}$:

$$M_t^a = \begin{cases} p_t^{-1}\left[p_0 M_0 + p_X \sum_{s=1}^{t}[X_s + a_s(M, X_1, \ldots, X_{s-1})]\right] & t \leq T \\ p_t^{-1}\left[p_0 M_0 + p_X \sum_{s=1}^{T}[X_s + a_s(M, X_1, \ldots, X_{s-1})] + p_X \sum_{s=T+1}^{t} X_s\right] & t > T \end{cases} \tag{24}$$

Proof. Henceforth, we let $a = (a_1, \ldots, a_T) \in \mathcal{A}$ denote the parent's optimal policy; also, to simplify the notation, we write M_{t-1}^a simply as M_{t-1}. Thus, by the arguments given in the text, at each time t, the child's optimal action is $b_t = M_{t-1}$.

Begin by analyzing the non-teaching periods. To this end, observe first that, from Eq. (24), for all $t \geq T$ and $\tau \geq 0$,

$$M_{t+\tau} = p_{t+\tau}^{-1}\left(p_t M_t + p_X \sum_{s=t+1}^{t+\tau} X_s\right), \tag{25}$$

where, as is customary, for $\tau = 0$, the empty summation is assumed to equal zero. Hence, for $t \geq T$ and $\tau \geq 1$,

$$\begin{aligned} X_{t+\tau} - M_{t+\tau-1} &= X_{t+\tau} - p_{t+\tau-1}^{-1} p_t M_t - p_{t+\tau-1}^{-1} p_X \sum_{s=t+1}^{t+\tau-1} X_s = \\ &= (X_{t+\tau} - M) - p_{t+\tau-1}^{-1} p_t (M_t - M) - p_{t+\tau-1}^{-1} p_X \sum_{s=t+1}^{t+\tau-1} (X_s - M); \end{aligned} \tag{26}$$

[36] There is evidence that agents are overconfident. See e.g. Fischoff, Slovic, and Lichtenstein [35], and Weinstein [70]. Recent papers explore the way in which overconfidence can be adaptive. See e.g. Compte and Postlewaite [29], Benabou and Tirole [9] and Koszegi [47].

the last line uses the fact that $p_t M + p_X \sum_{s=t+1}^{t+\tau-1} M = p_t M + p_X(\tau-1)M = p_{t+\tau-1}M$. It now follows that, at any time $t \geq T$, and for all $\tau \geq 1$, from the point of view of the parent, i.e. conditional upon the realization of M, the expected loss at time $t+\tau$ given the observed value of M_t is

$$E[(X_{t+\tau} - M_{t+\tau-1})'(X_{t+\tau} - M_{t+\tau-1})|M, M_t] =$$
$$= p_X^{-1} + (\tau-1)p_X p_{t+\tau-1}^{-1} p_X^{-1} p_{t+\tau-1}^{-1} p_X + (M_t - M)' p_t p_{t+\tau-1}^{-1} p_{t+\tau-1}^{-1} p_t (M_t - M);$$

this follows from the assumption that $X_1, \ldots, X_L$ are i.i.d. normal with mean M and precision p_X conditional on M,[37] which in turn implies that all of the cross-terms, which are of the form $(X_{t+\tau} - M)'(M_t - M)$, $(X_{t+\tau} - M)'(X_s - M)$ and $(M_t - M)'(X_s - M)$ for $s \in \{t+1, \ldots, t+\tau-1\}$, and $(X_s - M)'(X_\sigma - M)$ for s, σ distinct in $\{t+1, \ldots, t+\tau-1\}$, all have zero conditional expectation.

It follows that, for every non-teaching period $t+1 \in \{T+1, \ldots, L\}$, conditional upon M and M_t, the expected time-$(t+1)$ continuation value of the child's optimal policy is $V_{t+1}(M_t, M) = A_{t+1} + (M_t - M)' B_{t+1}(M_t - M)$, where[38]

$$A_{t+1} = \sum_{\tau=1}^{L-t} \delta^{\tau-1} \left[p_X^{-1} + (\tau-1) p_X p_{t+\tau-1}^{-1} p_X^{-1} p_{t+\tau-1}^{-1} p_X \right], \tag{27}$$

$$B_{t+1} = \sum_{\tau=1}^{L-t} \delta^{\tau-1} p_t p_{t+\tau-1}^{-1} p_{t+\tau-1}^{-1} p_t. \tag{28}$$

Turn now to teaching periods $t \in \{1, \ldots, T\}$. From the argument just given, $V_{T+1}(M_T, M) = A_T + (M_T - M)' B_{T+1}(M_T - M)$. We now show inductively that, for $t = T, \ldots, 1$, if $V_{t+1}(M_t, M) = A_{t+1} + (M_t - M)' B_{t+1}(M_t - M)$, then the equations for a_t and Γ_t in Thm. 4.2 hold, and furthermore $V_t(M_{t-1}, M) = A_t + (M_{t-1} - M)' B_t (M_{t-1} - M)$, where B_t is again as in Thm. 4.2.

Notice that, again from Eq. (24), for every $t \leq T$,

$$\begin{aligned} M_t &= p_t^{-1} p_0 M_0 + p_t^{-1} p_X \sum_{s=1}^{t} \Big(X_s + a_s(M, X_1, \ldots, X_{s-1}) \Big) = \\ &= p_t^{-1} p_{t-1} M_{t-1} + p_t^{-1} p_X \Big(X_t + a_t(M, X_1, \ldots, X_{t-1}) \Big). \end{aligned} \tag{29}$$

After substituting for M_t in the expression for V_{t+1} in the inductive hypothesis, conditional on the information $\mathcal{I}_t \equiv \{M, X_1, \ldots, X_{t-1}\}$, the action $a_t(M, X_1, \ldots, X_{t-1})$ must solve the Bellman equation

[37] Strictly speaking, since the parent knows M, in her view this is the actual unconditional distribution of the X_s's; for the child, M is a r.v., so the above statement describes the conditional distribution of the X_s's given M. However, we use a more explicit terminology and notation to remind the reader that M can be treated as a constant in the above expectations.

[38] For the purposes of solving for the optimal policy, we only need an expression for V_{T+1}. However, the expression for B_t, $t = T+1, \ldots, L$ is used below to show that $\gamma_t \in (0,1)$ in the scalar case.

$$V_t(M_{t-1}, M) = \min_{\bar{a}} \mathbf{E}\left[\bar{a}'C\bar{a} + (X_t + \bar{a} - M_{t-1})'(X_t + \bar{a} - M_{t-1})|\mathcal{I}_t\right] + \tag{30}$$
$$+\delta\mathbf{E}\left[A_{t+1} + \left(p_t^{-1}p_{t-1}M_{t-1} + p_t^{-1}p_X(X_t + \bar{a}) - M\right)' B_{t+1}\left(p_t^{-1}p_{t-1}M_{t-1} + p_t^{-1}p_X(X_t + \bar{a}) - M\right)|\mathcal{I}_t\right].$$

Differentiating with respect to $\bar{a}$, taking expectations, and dividing by 2 yields the FOC

$$\begin{aligned} 0 &= \bar{a}'C + \bar{a}' - (M_{t-1} - M)' + \\ &\quad +\delta\left(p_t^{-1}p_{t-1}M_{t-1} + p_t^{-1}p_X(M + a) - M\right)' B_{t+1}p_t^{-1}p_X \end{aligned}$$

(recall that C is symmetric) and therefore

$$\begin{aligned} a_t(M, X_1, \ldots, X_{t-1}) &= [C + I + \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_X]^{-1}[I - \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_{t-1}] \cdot (M_{t-1} - M) \\ &\equiv \Gamma_t(M_{t-1} - M). \end{aligned} \tag{31}$$

We now show that V_t can be expressed as a quadratic form in $(M_{t-1} - M)$. First, to simplify the quadratic form in the first line of Eq. (30), observe that

$$X_t + \Gamma_t(M_{t-1} - M) - M_{t-1} = (X_t - M) - (I - \Gamma_t)(M_{t-1} - M);$$

as for the quadratic form in the second line of Eq. (30),

$$\begin{aligned} &p_t^{-1}p_{t-1}M_{t-1} + p_t^{-1}p_X(X_t + \Gamma_t(M_{t-1} - M)) - M = \\ &= p_t^{-1}p_{t-1}(M_{t-1} - M) + p_t^{-1}p_X(X_t - M + \Gamma_t(M_{t-1} - M)) = \\ &= p_t^{-1}(p_{t-1} + p_X\Gamma_t)(M_{t-1} - M) + p_t^{-1}p_X(X_t - M). \end{aligned}$$

Now V_t can be rewritten as follows:

$$\begin{aligned} V_t(M_{t-1}, M) &= A_t + (M_{t-1} - M)'\Gamma_t'C\Gamma_t(M_{t-1} - M) + \\ &\quad + (M_{t-1} - M)'(I - \Gamma_t)'(I - \Gamma_t)(M_{t-1} - M) + \\ &\quad \delta(M_{t-1} - M)'(p_{t-1} + p_X\Gamma_t)'p_t^{-1}B_{t+1}p_t^{-1}(p_{t-1} + p_X\Gamma_t)(M_{T-1} - M) = \\ &\quad \equiv A_t + (M_{t-1} - M)'B_t(M_{t-1} - M), \end{aligned}$$

where A_t is a suitable constant, and the coefficient B_t in the last line are as in Thm. 4.2:

$$B_t = \Gamma_t'C\Gamma_t + (1 - \Gamma_t)'(1 - \Gamma_t) + \delta(p_{t-1} + p_X\Gamma_t)'p_t^{-1}B_{t+1}p_t^{-1}(p_{t-1} + p_X\Gamma_t). \tag{32}$$

■

Next, we derive certain useful expressions. First, notice that

$$\begin{aligned} I - \Gamma_t &= [C + I + \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_X]^{-1} \cdot \\ &\quad [C + I + \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_X - I + \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_{t-1}] = \\ &= [C + I + \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_X]^{-1}[C + \delta p_X p_t^{-1} B_{t+1}]. \end{aligned} \tag{33}$$

Similarly, in the scalar case,

$$\begin{aligned}\frac{p_{t-1}+\gamma_t p_X}{p_t} &= \frac{1}{p_t}\frac{p_{t-1}(1+C)+p_{t-1}\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2+p_X-p_X\delta B_{t+1}\frac{p_X p_{t-1}}{p_t^2}}{1+C+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2} = \\ &= \frac{1}{p_t}\frac{p_{t-1}(1+C)+p_X}{1+C+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2} = \\ &= \frac{1}{1+C+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2}\left(1+C\frac{p_{t-1}}{p_t}\right).\end{aligned} \tag{34}$$

Finally, again in the scalar case with $C = 0$, we can write B_t in the following useful form (for $t = 1, \ldots, T$):

$$\begin{aligned}B_t &= \left(\frac{\delta B_{t+1}\frac{p_X}{p_t}}{1+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2}\right)^2+\left(\frac{1}{1+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2}\right)^2\delta B_{t+1} = \\ &= \left(\frac{1}{1+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2}\right)^2\delta B_{t+1}\left(\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2+1\right) = \\ &= \frac{\delta B_{t+1}}{1+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2}\end{aligned} \tag{35}$$

which is the expression given in Thm. 2.1. For future reference, note also that

$$B_t = (1-\gamma_t)\frac{p_t}{p_X}. \tag{36}$$

Finally, we show that $\gamma_t \in (0, 1)$ for $t = 1, \ldots, T$ in the univariate case, as claimed in Thm. 2.1. It is clear that $\gamma_t < 1$, so we must only verify that $\gamma_t > 0$. Furthermore, notice that, if γ_t is positive for $C = 0$, then it is positive for all $C \geq 0$, and is decreasing in C. Thus, it is sufficient to prove the claim for $C = 0$.

Notice first that, even in the multivariate case, from Eq. (28), for $t = T+1, \ldots, L-1$,

$$\begin{aligned}
B_t &= \sum_{\tau=1}^{L-(t-1)} \delta^{\tau-1} p_{t-1} p_{t+\tau-2}^{-1} p_{t+\tau-2}^{-1} p_{t-1} = \\
&= p_{t-1} p_{t-1}^{-1} p_{t-1}^{-1} p_{t-1} + \sum_{\tau=2}^{L-(t-1)} \delta^{\tau-1} p_{t-1} p_{t+\tau-2}^{-1} p_{t+\tau-2}^{-1} p_{t-1} = \\
&= I + \delta \sum_{\tau=1}^{L-t} \delta^{\tau-1} p_{t-1} p_{t+\tau-1}^{-1} p_{t+\tau-1}^{-1} p_{t-1} = \\
&= I + \delta p_{t-1} p_t^{-1} \left(\sum_{\tau=1}^{L-t} \delta^{\tau-1} p_t p_{t+\tau-1}^{-1} p_{t+\tau-1}^{-1} p_t \right) p_t^{-1} p_{t-1} \\
&= I + \delta p_{t-1} p_t^{-1} B_{t+1} p_t^{-1} p_{t-1};
\end{aligned}$$

also, $B_L = I$. Furthermore, in the univariate case, we claim that $B_{t+1} \frac{p_X}{p_t} \frac{p_{t-1}}{p_t} < 1$ for all $t < L$. The claim is true for $t = L-1$, because $B_L = 1$ and $p_t = p_X + p_{t-1}$ for all $t \geq 1$. Now consider an arbitrary $t \in \{T+1, \ldots, L\}$ and assume the claim is true for $t+1, \ldots, L$. Now $B_t \frac{p_X}{p_{t-1}} \frac{p_{t-2}}{p_{t-1}} = \frac{p_X}{p_{t-1}} \frac{p_{t-2}}{p_{t-1}} + \delta B_{t+1} \left(\frac{p_{t-1}}{p_t} \right)^2 \frac{p_X}{p_{t-1}} \frac{p_{t-2}}{p_{t-1}} = \frac{p_X}{p_{t-1}} \frac{p_{t-2}}{p_{t-1}} + \delta B_{t+1} \frac{p_X}{p_t} \frac{p_{t-1}}{p_t} \cdot \frac{p_{t-1}}{p_{t-1}} \frac{p_{t-2}}{p_{t-1}} \leq \frac{p_X}{p_{t-1}} + \delta B_{t+1} \frac{p_X}{p_t} \frac{p_{t-1}}{p_t} \cdot \frac{p_{t-2}}{p_{t-1}} < 1$.

Hence, in particular, $\delta B_{T+1} \frac{p_X}{p_T} \frac{p_{T-1}}{p_T} < 1$, and so $\gamma_T \in (0,1)$. If $T = 1$, we are done. Otherwise, argue by induction; consider $t \in \{1, \ldots, T\}$ and assume that the claim is true for $t+1$. By Eq. (36), $B_{t+1} = (1-\gamma_{t+1}) \frac{p_{t+1}}{p_X}$, so $B_{t+1} \frac{p_X}{p_t} \frac{p_{t-1}}{p_t} = (1-\gamma_{t+1}) \frac{p_{t+1}}{p_X} \frac{p_X}{p_t} \frac{p_{t-1}}{p_t} = (1-\gamma_{t+1}) \frac{(p_t+p_X)(p_t-p_X)}{p_t^2} = (1-\gamma_{t+1}) \frac{p_t^2 - p_X^2}{p_t^2} = (1-\gamma_{t+1}) \left(1 - \frac{p_X^2}{p_t^2}\right) < (1-\gamma_{t+1})$, because $p_t = p_0 + t p_X \geq p_0 + p_X$ for $t \geq 1$. By the induction hypothesis, $1-\gamma_{t+1} \in (0,1)$, so $B_{t+1} \frac{p_X}{p_t} \frac{p_{t-1}}{p_t} < 1$ and therefore $\gamma_t \in (0,1)$, as claimed.

A.2 Other Results

A.2.1 Proof of Proposition 2.2

We have

$$\begin{aligned}
\gamma_t &= \frac{1 - \delta B_{t+1} \frac{p_X}{p_t} \frac{p_{t-1}}{p_t}}{1 + \delta B_{t+1} \left(\frac{p_X}{p_t} \right)^2} = \\
&= \frac{1 + \delta B_{t+1} \left(\frac{p_X}{p_t} \right)^2 \left(-\frac{p_{t-1}}{p_t} \right)}{1 + \delta B_{t+1} \left(\frac{p_X}{p_t} \right)^2} = \\
&= \mu_t \gamma_t^{IM} + (1-\mu_t) \gamma_t^{TL},
\end{aligned}$$

where the last equality follows from Eq. (11).

A.2.2 Proof of Proposition 2.3

For the first claim, recall that $\mu_t = \frac{1}{1+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2}$ and

$$B_{t+1}\left(\frac{p_X}{p_t}\right)^2 = \left(\frac{p_{t+1}}{p_t}\right)^2 \frac{\delta}{1+\delta B_{t+2}\left(\frac{p_X}{p_{t+1}}\right)^2} B_{t+2}\left(\frac{p_X}{p_{t+1}}\right)^2;$$

Clearly, if $\delta < \left(\frac{p_{T_0}}{p_{T_0+1}}\right)^2$, then $B_{T_0+1}\left(\frac{p_X}{p_{T_0}}\right)^2 < B_{T_0+2}\left(\frac{p_X}{p_{T_0+1}}\right)^2$ and so $\mu_{T_0} < \mu_{T_0+1}$; since $\frac{p_t}{p_{t+1}}$ is increasing in t, we have $\mu_t > \mu_{t+1}$ for all $t = T_0, \ldots, T$, as required.

For the second claim, recall from Eq. (32) that

$$B_{t+1} = (1-\gamma_{t+1})^2 + \delta B_{t+2}\left(\frac{p_t + \gamma_{t+1}p_X}{p_{t+1}}\right)^2.$$

Multiplying both sides by $\left(\frac{p_X}{p_t}\right)^2$ and expanding the square in the second term in the rhs yields

$$B_{t+1}\left(\frac{p_X}{p_t}\right)^2 = (1-\gamma_{t+1})^2\left(\frac{p_X}{p_t}\right)^2 + \delta B_{t+2}\left(\frac{p_X}{p_{t+1}}\right)^2 + \delta B_{t+2}\left(\frac{p_X}{p_t}\right)^2\left[\left(\frac{\gamma_{t+1}p_X}{p_{t+1}}\right)^2 + \left(\frac{2p_t\gamma_{t+1}p_X}{p_{t+1}}\right)\right].$$

Now observe that, by Eq. (11), $\gamma_{t+1} \in (0,1)$ also for $\delta = 1$. Hence, the above equation shows that, for $\delta = 1$, $B_{t+1}\left(\frac{p_X}{p_t}\right)^2 > B_{t+2}\left(\frac{p_X}{p_{t+1}}\right)^2$, and hence also $\mu_t < \mu_{t+1}$. Since both sides of the latter equation are continuous in δ, there is $\delta_t \in [0,1)$ such that $\delta > \delta_t$ implies $\mu_t < \mu_{t+1}$. Now take $\delta(T_0) = \min\{\delta_{T_0}, \delta_{T_0+1}, \ldots, \delta_T\}$ to obtain a cutoff with the required properties.

A.2.3 Proof of Proposition 2.5

Since $\mu_t = \frac{1}{1+\delta B_{t+1}\left(\frac{p_X}{p_t}\right)^2}$, it is enough to show that $B_{t+1}\left(\frac{p_X}{p_t}\right)^2$ is decreasing in p_0. For $t = T$, from Eq. (11) we have

$$B_{T+1}\left(\frac{p_X}{p_T}\right)^2 = p_X^2 \sum_{\tau=1}^{L-T} \delta^{\tau-1} \frac{1}{p_{T+\tau-1}^2}$$

which is clearly decreasing in p_0. Now, assuming the claim is true for $t+1 \leq T$, rewrite the equation used in the proof of the first claim in Prop. 2.3 as

$$B_{t+1}\left(\frac{p_X}{p_t}\right)^2 = \left(\frac{p_{t+1}}{p_t}\right)^2 \frac{\delta B_{t+2}\left(\frac{p_X}{p_{t+1}}\right)^2}{1+\delta B_{t+2}\left(\frac{p_X}{p_{t+1}}\right)^2}.$$

Now $\frac{p_{t+1}}{p_t} = \frac{p_0+(t+1)p_X}{p_0+tp_X}$ is strictly positive and decreasing in p_0;[39] furthermore, the fraction $\frac{x}{1+x}$ is increasing in x, and by the induction hypothesis $\delta B_{t+2}\left(\frac{p_X}{p_{t+1}}\right)^2$ is decreasing in p_0: thus, the second fraction in the rhs above is also strictly positive and decreasing in p_0. This completes the proof of the inductive step.

[39] The derivative of this fraction w.r.to p_0 is $p_t^{-2}[p_t - p_{t+1}] = -p_X p_t^{-2} < 0$.

A.2.4 Twin Experiment

To elaborate on the discussion in §3.3, we now describe an environment in which the parent and the child are ex-ante symmetric: just like the child, the parent is born (at time -1) with a prior Z_{-1}, drawn from the same distribution as M_{0a}; also, as we assume is the case for the child, before receiving additional information, the parent's prior over M is normal,with mean Z_{-1} and precision p_0. To reflect the fact that a child shares approximately 50% of a parent's genes, we assume that $\text{Corr}\left[Z_{-1}, M_{0a}\right] = \frac{1}{2}$.[40] As in Assumption 1, $\text{Cov}\left[Z_{-1}, X_0\right] = 0$. At time 0, before the child is born, the parent observes X_0.

Thus, Z_0 is now formally the parent's time-0 posterior mean, determined by her prior Z_{-1} and the observation of X_0. Under these assumptions, $Z_0 = \frac{p_0 Z_{-1} + p_X X_0}{p_0 + p_X}$. This implies that

$$\begin{aligned}\text{Cov}(M_{0a}, Z_0) &= \frac{p_0}{p_0 + p_X}\text{Cov}(M_{0a}, Z_{-1}) = \\ &= \frac{p_0}{p_0 + p_X}\frac{1}{2}\text{Var}\, M_{0a}\end{aligned}$$

because $\text{Corr}\left[M_{0a}, Z_{-1}\right] = \frac{1}{2}$ and M_{0a} has the same distribution (hence, the same variance) as Z_{-1}; similarly, since Z_{-1} is uncorrelated with X_0,

$$\text{Var}\, Z_0 = \left(\frac{p_0}{p_0 + p_X}\right)^2 \text{Var}\, M_{0a} + \left(\frac{p_X}{p_0 + p_X}\right)^2 \text{Var}\, X.$$

Therefore, the difference between the posterior covariances of twins raised together and twins raised apart, given by Eq. (17), can be rewritten as:

$$\left[\left(\frac{p_0}{p_0 + p_X}\right)^2 \frac{p_X \gamma_1}{p_0 + p_X} - \frac{p_0 + \gamma_1 p_X}{p_0 + p_X}\frac{p_0}{p_0 + p_X}\right]\text{Var}\, M_{0a} + \frac{p_X \gamma_1}{p_0 + p_X}\left(\frac{p_X}{p_0 + p_X}\right)^2 \text{Var}\, X.$$

The factor multiplying $\text{Var}\, M_{0a}$ can be simplified as

$$\begin{aligned}\frac{p_0}{p_0 + p_X}\left[\frac{p_0}{p_0 + p_X}\frac{p_X \gamma_1}{p_0 + p_X} - \frac{p_0 + \gamma_1 p_X}{p_0 + p_X}\right] &= \frac{p_0}{p_0 + p_X}\frac{p_0 p_X \gamma_1 - p_0^2 - p_0 p_X - p_0 p_X \gamma_1 - \gamma_1 p_X^2}{(p_0 + p_X)^2} = \\ &= \frac{p_0}{p_0 + p_X}\frac{-p_0^2 - p_0 p_X - \gamma_1 p_X^2}{(p_0 + p_X)^2}\end{aligned}$$

Therefore, recalling that $\text{Var}\, X = p_X^{-1}$, and canceling common terms, Eq. (17) has the same sign as

$$-p_0[p_0^2 + p_0 p_X + p_X^2 \gamma_1]\text{Var}\, M_{0a} + p_X^2 \gamma_1.$$

In particular, it is negative if and only if

$$\text{Var}\, M_{0a} > \frac{1}{p_0} \cdot \frac{p_X^2 \gamma_1}{p_0^2 + p_0 p_X + p_X^2 \gamma_1}.$$

A sufficient condition is thus $\text{Var}\left[M_{0a}\right] \geq \frac{1}{p_0}$.

[40] Strictly speaking, this incorporates the assumption that the effects of genetic endowment on an individual's ability to perform the task are linear. But this is consistent with similar assumptions that are prevalent in the behavioral-genetics literature.

A.2.5 Proof of Lemma 4.3

By Bayes' Rule, $f_{M|X_1+\gamma(M_0-M)}(m|x) \propto f_{X_1+\gamma(M_0-M)|M}(x|m)f_M(m)$, where the f's denote the relevant conditional and prior densities and "$\propto$" is the proportionality symbol (i.e. the l.h.s. equals the r.h.s. times a normalizing factor that does not depend on m). Recall that $M \sim \mathrm{N}(M_0, \frac{1}{p_0})$, and note that

$$X_1 + \gamma(M_0 - M)|M \sim \mathrm{N}\left(\gamma M_0 + (1-\gamma)M, \frac{1}{p_X}\right);$$

therefore, $f_{X_1+\gamma(M_0-M)|M}(x|m)f_M(m) \propto e^{-\frac{1}{2}Q(x,m)}$, where

$$\begin{aligned}
Q(x,m) &= p_X[x - \gamma M_0 - (1-\gamma)m]^2 + p_0(m-M_0)^2 = \\
&= p_X[(x-M_0) - (1-\gamma)(m-M_0)]^2 + p_0(m-M_0)^2 = \\
&= p_X(1-\gamma)^2\left[\frac{x-M_0}{1-\gamma} - (m-M_0)\right]^2 + p_0(m-M_0)^2 = \\
&= p_X(1-\gamma)^2\left(\frac{x-M_0}{1-\gamma}\right)^2 + [p_0 + p_X(1-\gamma)^2](m-M_0)^2 - \\
&\quad -2p_X(1-\gamma)^2\frac{x-M_0}{1-\gamma}(m-M_0) = \\
&= [p_0 + p_X(1-\gamma)^2]\left[\frac{p_X(1-\gamma)^2}{p_0 + p_X(1-\gamma)^2}\frac{x-M_0}{1-\gamma} - (m-M_0)\right]^2 + K(x) = \\
&= [p_0 + p_X(1-\gamma)^2]\left[M_0 + \frac{p_X(1-\gamma)}{p_0 + p_X(1-\gamma)^2}(x-M_0) - m\right]^2 + K(x) = \\
&= [p_0 + p_X(1-\gamma)^2]\left[\frac{p_0 - p_X\gamma(1-\gamma)}{p_0 + p_X(1-\gamma)^2}M_0 + \frac{p_X(1-\gamma)}{p_0 + p_X(1-\gamma)^2}x - m\right]^2 + K(x),
\end{aligned}$$

where $K(x) = p_X(1-\gamma)^2\left(\frac{x-M_0}{1-\gamma}\right)^2 - \frac{p_X^2(1-\gamma)^4}{p_0+p_X(1-\gamma)^2}\left(\frac{x-M_0}{1-\gamma}\right)^2$ is the constant required to "complete the square." By inspecting $Q(x,m) - K(x)$, it follows that $f_{M|X_1+\gamma(M_0-M)}$ must be a normal density with mean

$$\frac{p_0 - p_X\gamma(1-\gamma)}{p_0 + p_X(1-\gamma)^2}M_0 + \frac{p_X(1-\gamma)}{p_0 + p_X(1-\gamma)^2}[X_1 + \gamma(M_0 - M)]$$

and precision $p_0 + p_X(1-\gamma)^2$.

A.2.6 Proof of Proposition 4.4 and Corollary 4.5.

Begin by writing the parent's objective function in Eq. (23) in a way that is consistent with both the model of Sec. 2.1 and Sec. 4.3. In both cases, the child's guess at time 2 takes the form $w(X_1 + \bar{a}) + (1-w)M_0$: if the child is unaware of the parent's action, then $w = \frac{p_X}{p_0+p_X}$, whereas, under the textbook equilibrium assumption, Lemma 4.3 implies that $w = \frac{p_X(1-\gamma^e)}{p_0+p_X(1-\gamma^e)^2}$. We get

$$\mathrm{E}\left[(X_1 + \bar{a} - M_0)^2 + \delta[X_2 - (1-w)M_0 - w(X_1+\bar{a})]^2|M\right]. \tag{37}$$

It is also convenient to calculate the expectation: adding and subtracting M inside the two squared-loss terms, we get

$$\begin{aligned}
&\mathrm{E}\Big[(X_1+\bar{a}-M)+(M-M_0)]^2+ \qquad (38)\\
&+\delta[(X_2-M)+(1-w)(M-M_0)+w(M-X_1-\bar{a})]^2|M\Big]=\\
&=\bar{a}^2+\frac{1}{p_X}+(M_0-M)^2+2\bar{a}(M-M_0)+\\
&+\delta\left\{\frac{1}{p_X}+(1-w)^2(M_0-M)^2+w^2\left(\frac{1}{p_X}+\bar{a}^2\right)+2w(1-w)(M_0-M)\bar{a}\right\}=\\
&=(1+\delta w^2)\bar{a}^2-2[1-\delta w(1-w)](M_0-M)\bar{a}+ \qquad (39)\\
&+\delta\left\{(1-w)^2(M_0-M)^2+w^2\frac{1}{p_X}\right\}+\frac{1+\delta}{p_X}+(M_0-M)^2.
\end{aligned}$$

Differentiating with respect to $\bar{a}$ and rearranging terms yields

$$\bar{a}=\frac{1-\delta(1-w)w}{1+\delta w^2}(M_0-M), \qquad (40)$$

and it is also immediate to verify that the second derivative of Eq. (39) is strictly positive.

For $w=\frac{p_X}{p_0+p_X}$, we obtain the optimal intensity $\gamma_T=\gamma_1$ in Eq. (11) of Thm. 2.1 (where $B_{T+1}=B_L=B_2=1$).

In the textbook-equilibrium case, Eq. (40) instead implies that a necessary and sufficient condition for a linear equilibrium with intensity γ^e is $\gamma^e=\frac{1-\delta(1-w)w}{1+\delta w^2}$. This can be rewritten as follows: notice that $[p_0+p_X(1-\gamma^e)^2]^2w(1-w)=p_X(1-\gamma^e)[p_0-p_X\gamma^e(1-\gamma^e)]$; then the required condition is

$$\gamma^e=\frac{[p_0+p_X(1-\gamma^e)^2]^2-\delta p_X(1-\gamma^e)[p_0-p_X\gamma^e(1-\gamma^e)]}{[p_0+p_X(1-\gamma^e)^2]^2+\delta p_X^2(1-\gamma^e)^2}.$$

Multiplying by the denominator, which is strictly positive, and rearranging yields

$$[p_0+p_X(1-\gamma^e)^2]^2(1-\gamma^e)-\delta p_X(1-\gamma^e)[p_0-p_X\gamma^e(1-\gamma^e)+\gamma^e p_X(1-\gamma^e)]=0$$

which can be rewritten as

$$(1-\gamma^e)\Big\{[p_0+p_X(1-\gamma^e)^2]^2-\delta p_0 p_X\Big\}=0.$$

Therefore, one solution is always $\gamma^e=1$. To find other solutions, assume $\gamma^e\neq 1$, divide by $1-\gamma^e$ and rearrange:

$$p_X(1-\gamma^e)^2=\sqrt{\delta p_0 p_X}-p_0.$$

Since $p_0>0$, the rhs is positive provided $\delta p_X>p_0$. If this condition holds, then the above holds if and only if

$$1-\gamma^e=\pm\sqrt{\frac{\sqrt{\delta p_0 p_X}-p_0}{p_X}}.$$

Note also that, if $\delta p_X > p_0$, then $\sqrt{\delta p_0 p_X} - p_0 < \delta p_X - p_0 < p_X$, which implies that the argument of the outer root is strictly between 0 and 1.

Choosing the positive sign in the above equation yields the expression for γ^e in Prop. 4.4, henceforth denoted γ^+; by the argument just given, $\gamma^+ \in (0,1)$; Choosing the negative sign instead yields $\gamma^- \equiv 2 - \gamma^+$.

To complete the proof of Cor. 4.5, it is convenient to express the minimized value of the parent's objective function by substituting for $\bar{a}$ in Eq. (39) using Eq. (40); we get

$$\begin{aligned}
&\frac{[1-\delta(1-w)w]^2}{1+\delta w^2}(M_0-M)^2 - 2\frac{[1-\delta w(1-w)]^2}{1+\delta w^2}(M_0-M)^2 + \\
&+\delta\left\{(1-w)^2(M_0-M)^2 + w^2\frac{1}{p_X}\right\} + \frac{1+\delta}{p_X} + (M_0-M)^2 = \\
&=\frac{-[1-\delta w(1-w)]^2 + \delta(1-w)^2 + \delta^2(1-w)^2w^2 + 1 + \delta w^2}{1+\delta w^2}(M_0-M)^2 + \frac{1+\delta(1+w^2)}{p_X} = \\
&=\frac{2\delta w(1-w) + \delta(1-w)^2 + \delta w^2}{1+\delta w^2}(M_0-M)^2 + \frac{1+\delta(1+w^2)}{p_X} = \\
&=\frac{\delta(1-w)[2w+1-w] + \delta w^2}{1+\delta w^2}(M_0-M)^2 + \frac{1+\delta(1+w^2)}{p_X} = \\
&=\frac{\delta}{1+\delta w^2}(M_0-M)^2 + \frac{1+\delta(1+w^2)}{p_X}; \qquad (41)
\end{aligned}$$

since $w^2 = \frac{p_X^2(1-\gamma^e)^2}{[p_0+p_X(1-\gamma^e)^2]^2}$ and $(1-\gamma^-)^2 = (\gamma^+-1)^2 = (1-\gamma^+)^2$, it follows that the payoff to the parent in the two equilibria with intensity coefficients γ^+ and γ^- is the same.

Finally, suppose that $\bar{a} = \gamma^e(M_0 - M)$ is an equlibrium parental intervention, with $b_1 = M_0$ and b_2 as in Lemma 4.3, and fix $k \in \mathbb{R}$. Then let $\bar{a}' = \bar{a} + k$, $b_1' = b_1 + k$ and $b_2' = b_2$; clearly, (a, b) constitutes an equilibrium in which the child "filters out" the additive constant k both in her time-1 guess and in her updating rule. It is immediate to verify that this new equilibrium is payoff-equivalent to the original one.

References

[1] Apesteguia, Jose, Steffen, and Oechssler, Jrg(2006): "Imitation: theory and experimental evidence" *Journal of Economic Theory*, forthcoming.

[2] Altmann (1980) *Baboon Mothers and Infants* The University of Chicago Press.

[3] Bandura A. (1977): *Social Learning Theory*, Prentice-Hall.

[4] Bandura A. and R.H. Walters (1963): *Social Learning and Personality Development*, New York: Holt, Rinehart & Winston.

[5] Baumrind, D. (1967): "Child care practices anteceding three patterns of preschool behavior," *Genetic Psychology Monographs*, 75, 43-88.

[6] Baydar, N. and Brooks-Gunn, J. (1991): "Effects of maternal employment and childcare arrangements on preschoolers cognitive and behavioral outcomes: evidence from the Children of the NLSY." *Developmental Psychology*, vol. 27, 932-45.

[7] Becker, G. S. (1981): *A Treatise on the Family*. Harvard University Press, Cambridge, Massachusetts.

[8] Belsky, J. and Eggebeen, D. (1991): "Early and extensive maternal employment and young children's socioemotional development: Children of the National Longitudinal Survey of Youth." *Journal of Marriage and the Family*, vol. 53, 1083-1110.

[9] Benabou, Roland and Jean Tirole (2002) "Self-Confidence and Personal Motivation" *Quarterly Journal of Economics,* 117(3), 871-915.

[10] Benabou, Roland and Jean Tirole (2003) "Intrinsic and Extrinsic Motivation" *Review of Economic Studies,* 70(3) (2003), 489-520.

[11] Berardo, J., and Smith, A. (2000): *Bayesian Theory*, 2nd ed., John Wiley and Sons, New York.

[12] Bernal, Raquel (2004): "Employment and Child Care Decisions of Mother's and the Well-being of their Children," manuscript, Northwestern University.

[13] Bisin, A. and T. Verdier (2000): "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits," *Quarterly Journal of Economics*, CXV(3), 955-988.

[14] Bisin, A. and T. Verdier (2001): "Cultural Transmission, Socialization, and the Dynamics of Preferences," *Journal of Economic Theory*, 97(1), 298-319.

[15] Bisin, A., G. Topa and T. Verdier (2004): "Religious Intermarriage and Socialization in the U.S.", *Journal of Political Economy*, 112-3, 615-64.

[16] Bjorklund, Anders, Lindahl, Michael and Erik Plug (2006): "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *Quarterly Journal of Economics*, vol. CXXI n. 3, pp. 999-1029.

[17] Black, Sandra, Devereux, Paul, and Salvanes, Kjell (2005): "The More the Merrier? The Effect of Family Size and Birth Order on Childrens Education," *Quarterly Journal of Economics*, Vol. 120, N. 2 669-700.

[18] Blurton-Jones, Nicholas: "The lives of hunter - gatherer children" in *Juvenile primates : life history, development, and behavior* edited by Michael E. Pereira, Lynn A. Fairbanks; New York : Oxford University Press, 1993.

[19] Bouchard, T.J. Jr., Lykken, D.T., McGue, M., Segal, N.L., and Tellegen, A. (1990): "Sources of Human Psychological Differences: The Minnesota Study of Twins Reared apart," *Science*, 250, 223-338.

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

[21] Brezzi, M. and Lai, T.L. (2000): "Incomplete Learning from Endogenous Data in Dynamic Allocation," *Econometrica* vol. 68 n. 6, 1511-1516.

[22] Brunnermeier, Markus and Jonathan Parker (2004). "Optimal Expectations," Manuscript, Princeton University.

[23] Camerer, Colin F. (2003): *Behavioral Game Theory: Experiments in Strategic Interaction*; Princeton University Press.

[24] Carneiro, Pedro, Heckman, James J. and Dimitriy V. Masterov (2002): "Labor Market Discrimination and Racial Differences in Premarket Factors," manuscript, University of Chicago.

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

[26] Chisholm, K. "A Three Year Follow-up of Attachment and Indiscriminate Friendliness in Children Adopted from Romanian Families." *Child Developmen* 69(4): 1092-1106.

[27] Clutton-Brock, T. H. (1991): *The evolution of parental care* Princeton, N.J. : Princeton University Press.

[28] Collins, W.A., Maccoby, E., Steinberg, L., Hetherington, E.M., and Bornstein, M.H. (2000): "Contemporary research on parenting: the case for nature and nurture," *American Psychologist*, **55** (2), 218-232.

[29] Compte, Olivier and Andrew Postlewaite: "Confidence-Enhanced Performance," *American Economic Review*, forthcoming.

[30] Crawford, Vincent and Joel Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50 n. 6, pp. 1431-1451.

[31] Demo, D. H., and Cox, M. J. (2000): "Families with young children: A review of research in the 1990s," *Journal of Marriage and the Family*, 62, 867-895.

[32] Desai, S., Chase-Lansdale, P. L. and Michael, R. T. (1989): "Mother or market? effects of maternal employment on the intellectual ability of 4-year old children." *Demography*, vol. 26, 545-61.

[33] Easley, David and Nicholas Kiefer "Controlling a Stochastic Process with Unknown Parameters;" *Econometrica*, 1988.

[34] Ettinger, David and Philippe Jehiel (2006): "Towards a Theory of Deception," manuscript, PSE and University College London.

[35] Fischoff, B., Slovic, P., and Lichtenstein, S. (1977): "Knowing with Certainty: The Appropriateness of Extreme Confidence," *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), 552-564.

[36] Fudenberg Drew and David K. Levine (1998): *Theory of Learning in Games*; MIT Press.

[37] Gilboa, I., and D. Schmeidler, "Case-Based Decision Theory", *The Quarterly Journal of Economics*, **110** (1995) 605-639

[38] Gopnick, A., and Meltzoff, A. (1992): "Categorization and naming: basic-level sorting in eighteen-month-olds and its relationship to language," *Child Development*, **63**, 1091-1103.

[39] Griffin, Gary A. and Harry F. Harlow (1966) "Effects of Three Months of Total Social Deprivation on Social Adjustment and Learning in the Rhesus Monkey," *Child Development,* Vol. 37, No. 3. pp. 533-547.

[40] Harlow, H. F., and Zimmerman, R. (1959). "Affectional responses in the infant monkey." *Science,* 130, 421-432.

[41] Harbaugh, W., K. Krause, and T. Berry (2001) "AGARP for Kids: On the Development of Rational Choice Behavior," *American Economic Review*, 91 #5, Dec., 1539-1545.

[42] Harris, Judith Rich (1995): "Where Is the Child's Environment? A Group Socialization Theory of Development," *Psychological Review*, Vol. 102, N. 3, 458-489.

[43] Harris, Judith Rich (1998): *The Nurture Assumption.* The Free Press.

[44] Hoopes, Janet L. (1982): *Prediction in Child Development: A Longitudinal Study of Adoptive and Non-Adoptive Families. The Delaware Family Study.* Child Welfare Leagues of America, New York, NY.

[45] Harsanyi, J. (1967-68): "Games of Incomplete Information Played by Bayesian Players. Parts I, II, III," *Management Science*, 14, 159-182, 320-334, 486-502.

[46] Kalai, E., and Lehrer, E. (1993): "Rational learning leads to Nash equilibrium," *Econometrica* **61** (5), 1019-1045.

[47] Koszegi, Botond (2000) "Ego Utility, Overconfidence, and Task Choice," manuscript, University of California, Berkeley.

[48] Lamport-Commons, M., and Miller, P. (1992): "Emotional learning in infants: a cross-cultural examination,"*Developmental Psychology*, **28** n.4.

[49] Lancaster Jane B., and Chet S. Lancaster (1987): "The Watershed: Change in Parental-Investment and Family Formation Strategies in the Course of Human Evolution." in *Parenting Across the Lifespan* edited by Jane B. Lancaster, Jeanne Altmann, Alice S. Rossi, And Lonnie Sherrod; Aldyne de Gruyter, Hawthorne, NY.

[50] OECD Publishing (2006): "Education at a Glance: OECD Indicators - 2005 Edition."

[51] Pinker, S. (2002): *The Blank Slate: The Modern Denial of Human Nature*, Viking-Penguin.

[52] Plomin, R. and Daniels, D. (1987): "Why are children in the same family so different from each other?" *The Behavioral and Brain Sciences,* 10, 1-16.

[53] Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001): *Behavioral Genetics.* (4th Edition). New York: Worth Publishers.

[54] Reiss David, Neiderheiser Jenae M., Hetherington Mavis E. and Plomin Robert (2000): *The Relationship Code.* Harvard University Press.

[55] Richman, A.L., LeVine, R., Stables, R., Howrigan, G., Welles-Nystrom, B., and LeVine, S. (1988): "Maternal behavior to infants in five cultures,"in *Parental Behavior in Diverse Societies*, LeVine, R., Miller, P., and West, M. (eds.), Jossey-Bass Inc., San Francisco

[56] Ridley, Matt, (2003): *Nature Via Nurture : Genes, Experience, and What Makes Us Human*, Harper Collins.

[57] Robson, Arthur, Kaplan, Hillard, Lancaster, Jane (2006): "Embodied Capital and the Evolutionary Economics Of the Human Lifespan," *Population and Development Review*, forthcoming.

[58] Rogoff, B. (1997): "Cognition as a collaborative process," in *Handbook of Child Psychology, Vol. 2*, Damon, W. (ed.), fifth edition, John Wiley and Sons, New York.

[59] Rubinstein, Ariel (1989) "Optimal Flexibility of Rules: The Tale of the Wise Principal and the Naive Agent," *European Journal of Political Economy* 5 219-227.

[60] Sacerdote, Bruce (2002): "The Nature and Nurture of Economic Outcomes," *American Economic Review Papers and Proceedings*, vol. XCII p. 344–348.

[61] Sacerdote, Bruce (2006): "How Large are the Effects from Changes in the Family Environment? A Study of Korean-American Adoptees," *Quarterly Journal of Economics*, forthcoming.

[62] Shonkoff, Jack, and Deborah Phillips, eds. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development.* Washington, D.C.: National Academy Press.

[63] Smith, Lones and Peter Sorensen (2006): "Informational Herding and Optimal Experimentation," manuscript, University of Michigan.

[64] Schlag, Karl H. (1998) "Why Imitate, and If So, How?, : A Boundedly Rational Approach to Multi-armed Bandits" *Journal of Economic Theory*, Volume 78, Issue 1, Pages 130-156

[65] Todd, Petra E., Kenneth I. Wolpin, (2004). "The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps" manuscript, University of Pennsylvania.

[66] Turkheimer, Eric, Haley, Andreanna, Waldron, Mary, D'Onofrio, Brian, and Gottesman, Irving I. (2003): "Socioeconomic Status Modifies Heritability of IQ in Young Children," *Psycological Science* XIV 623-628

[67] Vandell D.L. and J. Ramanan (1992) "Effects of Early and Recent Maternal Employment on Children from Low Income Families." *Child Development* 63: 938-949.

[68] Vygotsky L.S. (1978). *Mind in Society.* Cambridge, MA: Harvard University Press.

[69] Waxman, S. (in press): "Links between object categorization and naming: Origins and emergence in human infants," In D. K. Rakison & L. M. Oakes (Eds.), *Early category and concept development: Making sense of the blooming, buzzing confusion.* New York: Oxford University Press.

[70] Weinstein, N . D . (1980): "Unrealistic Optimism About Future Life Events," *Journal of Personality and Social Psychology,* 39(5), 806—820.

[71] Warren, S. (1992): "Lower threshold for referral for psychiatric treatment for adopted adolescents." *Journal of the American Academy of Child and Adolescent Psychiatry* 31 n. 3.

[72] Wood, D., Bruner, J.S., and Ross, G. (1976): "The role of tutoring in problem-solving," *Journal of Child Psychology and Psychiatry*, **17**, 89-100.

[73] Wood, D., and Middleton, D. (1975): "A study of assisted problem-solving," *British Journal of Psychology*, **66**, 181-191.

[74] Wood, D., Wood, H., and Middleton, D. (1978): "An experimental evaluation of four face-to-face teaching strategies," *International Journal of Behavioral Development*, **2**, 131-147.

[75] Zentall, Thomas R. and Bennet G. Galef JR. (1988) *Social Learning: Psychological and Biological Perspectives.* Lawrence Erlbaum Associates, Publishers.

B Online Appendix

B.0.7 Proof of Theorem 3.1 (sketch)

The claim essentially follows from the Certainty Equivalence Principle. For completeness, we briefly sketch how to modify the proof of Theorem 4.2 given above for the case in which the parent's information set at the beginning of time t is $\mathcal{I}_t = \{X_1, \ldots, X_{t-1}\}$ (notice that this yields a multivariate generalization of Theorem 3.1 that also includes a cost term). Loosely speaking, the basic idea is to replace each occurrence of M at time t with the parent's posterior mean, $Z_t = p_{Zt}^{-1}[p_{Z0}Z_0 + p_X \sum_{s=1}^{t-1} X_s]$, where $p_{Zt} = p_{Z0} + tp_X$.

First, analogously to Eq. (26), for $t \geq T$ and $\tau \geq 1$ we can write

$$X_{t+\tau} - M_{t+\tau-1} = (X_{t+\tau} - Z_t) - p_{t+\tau-1}^{-1} p_t (M_t - Z_t) - p_{t+\tau-1}^{-1} p_X \sum_{s=t+1}^{t+\tau-1} (X_s - Z_t) :$$

this follows directly from the decomposition in Eq. (25). In turn, this implies that

$$\begin{aligned} &E[(X_{t+\tau} - M_{t+\tau-1})'(X_{t+\tau} - M_{t+\tau-1})|\mathcal{I}_{t+1}] = \\ &= \text{const.} + (M_t - Z_t)' p_t p_{t+\tau-1}^{-1} p_{t+\tau-1}^{-1} p_t (M_t - Z_t); \end{aligned}$$

note that the constant term is different here, because the conditional variance of X_s, for $s > t$, given Z_t is not p_X, and the X_s's are no longer independent conditional upon Z_t. However, it is still the case that the only non-zero term involving M_t is the quadratic form in $(M_t - Z_t)$: for $s = t+1, \ldots, t+\tau$, $E[(X_s - Z_t)'(M_t - Z_t)|\mathcal{I}_{t+1}] = E[X_s - Z_t|\mathcal{I}_{t+1}]'(M_t - Z_t) = 0$, because Z_t is the parent's estimate of the mean of X_s at the beginning of time $t+1$. As above, for every non-teaching period $t+1 \in \{T+1, \ldots, L\}$, we can then write $V_{t+1}(M_t, Z_t) = A_{t+1} + (M_t - Z_t)' B_{t+1}(M_t - Z_t)$, where A_{t+1} is suitable constant and B_{t+1} is exactly as in Theorem 4.1.

Turn now to teaching periods $t \in \{1, \ldots, T\}$. Again, $V_{T+1}(M_T, M) = A_T + (M_T - M)' B_{T+1}(M_T - M)$. The proof of the inductive step is as in Theorem 2.1; the optimal action $a_t(X_1, \ldots, X_{t-1})$ must now solve the Bellman equation

$$\begin{aligned} V_t(M_{t-1}, Z_{t-1}) = \min_{\bar{a}} \mathbf{E}\left[\bar{a}' C \bar{a} + (X_t + \bar{a} - M_{t-1})'(X_t + \bar{a} - M_{t-1})|\mathcal{I}_t\right] + \\ +\delta \mathbf{E}\left[A_{t+1} + \left(p_t^{-1} p_{t-1} M_{t-1} + p_t^{-1} p_X (X_t + \bar{a}) - Z_t\right)' B_{t+1} \left(p_t^{-1} p_{t-1} M_{t-1} + p_t^{-1} p_X (X_t + \bar{a}) - Z_t\right)|\mathcal{I}_t\right]. \end{aligned} \tag{42}$$

Differentiating with respect to $\bar{a}$, taking expectations, and dividing by 2 yields the FOC

$$\begin{aligned} 0 \;=\; &\bar{a}' C + \bar{a}' - (M_{t-1} - Z_{t-1})' + \\ &+\delta \left(p_t^{-1} p_{t-1} M_{t-1} + p_t^{-1} p_X (Z_{t-1} + \bar{a}) - Z_{t-1}\right)' B_{t+1} p_t^{-1} p_X, \end{aligned}$$

where we use the fact that $E[X_{t+1}|\mathcal{I}_t] = E[Z_t|\mathcal{I}_t] = Z_{t-1}$. Therefore,

$$\begin{aligned} a_t(X_1, \ldots, X_{t-1}) &= [C + I + \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_X]^{-1} [I - \delta p_X p_t^{-1} B_{t+1} p_t^{-1} p_{t-1}] \cdot (M_{t-1} - Z_{t-1}) \\ &= \Gamma_t (M_{t-1} - Z_{t-1}), \end{aligned} \tag{43}$$

where Γ_t is as in Theorem 4.1. To show that V_t has the required quadratic form, we note that $X_t + \Gamma_t(M_{t-1} - Z_{t-1}) - M_{t-1} = (X_t - Z_{t-1}) - (I - \Gamma_t)(M_{t-1} - Z_{t-1})$, as in the original proof; furthermore,

$$\begin{gathered} p_t^{-1} p_{t-1} M_{t-1} + p_t^{-1} p_X (X_t + \Gamma_t (M_{t-1} - Z_{t-1})) - Z_t = \\ = p_t^{-1} p_{t-1} M_{t-1} + p_t^{-1} p_X (X_t + \Gamma_t (M_{t-1} - Z_{t-1})) - Z_{t-1} + (Z_{t-1} - Z_t) = \\ = p_t^{-1} p_{t-1} (M_{t-1} - Z_{t-1}) + p_t^{-1} p_X (X_t - Z_{t-1} + \Gamma_t (M_{t-1} - Z_{t-1})) + (Z_{t-1} - Z_t) = \\ = p_t^{-1} (p_{t-1} + p_X \Gamma_t) (M_{t-1} - Z_{t-1}) + p_t^{-1} p_X (X_t - Z_{t-1}) + (Z_{t-1} - Z_t). \end{gathered}$$

The key observation is now that the expectation of the quadratic form in the second line of Eq. (42) reduces to a constant plus

$$(M_{t-1} - Z_{t-1})' (p_{t-1} + p_X \Gamma_t)' p_t^{-1} B_{t+1} p_t^{-1} (p_{t-1} + p_X \Gamma_t) (M_{t-1} - Z_{t-1}),$$

which implies that $V_t(M_{t-1}, Z_{t-1}) = A_t + (M_{t-1} - Z_{t-1})' B_t (M_{t-1} - Z_{t-1})$, with B_t as in Thm. 4.1 (the constant A_t however is different). Again, the key observation is that all other cross-terms involving M_{t-1} have zero expectation conditional on $\mathcal{I}_t$.