ESSAYS IN ECONOMICS

by

Xiangrong Yu

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Economics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2012

Date of final oral examination: May 11, 2012

The dissertation is approved by the following members of the Final Oral Committee:
Steven Durlauf, Professor, Economics
William Brock, Professor, Economics
Kenneth West, Professor, Economics
Chao Fu, Assistant Professor, Economics
Laura Schechter, Associate Professor, Agricultural and Applied Economics
ACKNOWLEDGMENTS

This dissertation largely concludes my Ph.D. studies in Madison. During these five years, I have been fortunate to meet and work with many outstanding people.

I would like to express my gratitude to Steven Durlauf for his guidance, support, and friendship. As my principal advisor, Steven has provided me with an excellent atmosphere for doing research. His influence pervades my work. Steven is also a great mentor in many other aspects, from professional development to academic writing to even understanding of American culture. I believe he will continue to be an important source of encouragement and inspiration in my life after graduation. He always said “our intellectual interaction is a lifetime contract.” And I will always be in his debt.

I would also like to thank Buz Brock and Ken West for their great advice and continuous support. Buz is the only professor from whom I have taken three courses. I have learned a great deal of techniques and ideas from numerous communications with him. Ken has always been willing to help and give his best suggestions, and has shown great patience with me. I have benefited from his invaluable feedback and comments on my papers.

I was fortunate to receive advice from and be inspired by Gary Becker when I was visiting the University of Chicago. Many conversations with him have taught me the art of asking relevant questions and telling economic stories. I wish I could gain even one percent of his skills of economic analysis. His comments on my work and support for my job search are greatly acknowledged.
I want to thank Chao Fu and Laura Schechter for their generous help. I am also grateful to Wallice Ao, John Cochrane, Dean Corbae, Kenichi Fukushima, Chao He, Shuanglin Lin, Dan Quint, Bill Sandholm, Shane Su, Scott Swisher, Chris Taber, Man-Keung Tang, Noah Williams, Yang Yao, and many friends for their insights and stimulating discussions at different stages of my Ph.D. work.

My appreciation to my wife and daughter, parents, and sister is beyond words.
CONTENTS

Acknowledgements  i

Contents  iii

1 Introduction  1

2 Family Linkages and Social Interactions in Human Capital Formation  4
  2.1 Introduction  5
  2.2 Parent-Child Interactions in the Beckerian Model of Human Capital  13
  2.3 Parental Investments and Sibling Interactions  33
  2.4 Parental Investments and Neighborhood Interactions  55
  2.5 Conclusion  70

3 Measurement Error and Policy Evaluation in the Frequency Domain  73
  3.1 Introduction  74
  3.2 General Framework  82
  3.3 Robust Policy Rule 102
  3.4 Some Applications: Monetary Policy Evaluation 107
  3.5 Concluding Remarks 131

A Appendix of Chapter 2: Derivations and Proofs 134

B Appendix of Chapter 3: Some Examples and Results 152

References 159
1 INTRODUCTION
This thesis consists of two self-contained essays in economics, which form the second and third chapters, respectively.

The first essay introduces parent-child interactions into the Beckerian model of human capital. The acquisition of human capital, jointly determined by parental investment and child effort, is an equilibrium outcome of the intergenerational interactions, which is Pareto efficient within the family. This essay shows that the equilibrium output of human capital is not affected by the parental authority over child behavior, which in the spirit is analogous to Becker’s (1974) "Rotten Kid Theorem," but is usually lower than the level that maximizes the instantaneous aggregate family welfare. In a family with more than one child, siblings not only compete for parental investments but also directly interact with each other in their effort choices. Exploring these intragenerational connections and their interplay with intergenerational forces, this essay presents a more complete theory of family linkages in the development of human capital and the associated implications for the rise and fall of families. Social interactions among children from different families induce intragenerational feedback effects that are further amplified by intrafamily interactions and accelerate regression towards the mean in the economic status of families. Allowing for endogenous group formation, this essay also characterizes the conditions, which reflect the interplay between family influences and social effects, for the emergence of segregation in an equilibrium. Under a linear quadratic specification studied throughout this essay, the notion of social multipliers is generalized to conceptualize the role of parent-child interactions in peer behaviors.
The second essay explores frequency-specific implications of measurement error for the design of stabilization policy rules. Policy evaluation in the frequency domain is interesting because the characterization of policy effects frequency by frequency gives the policymaker additional information about the effects of a given policy. Further, some important aspects of policy analysis can be better understood in the frequency domain than in the time domain. In this essay, I develop a rich set of design limits that describe fundamental restrictions on how a policymaker can alter variance at different frequencies. I also examine the interaction of measurement error and model uncertainty to understand the effects of different sources of informational limit on optimal policymaking. In a linear feedback model with noisy state observations, measurement error seriously distorts the performance of the policy rule that is optimal for the noise-free system. Adjusting the policy to appropriately account for measurement error means that the policymaker becomes less responsive to the raw data. For a parameterized example which corresponds to the choice of monetary policy rules in a simple AR (1) environment, I show that an additive white noise process of measurement error has little impact at low frequencies but induces less active control at high frequencies, and even may lead to more aggressive policy actions at medium frequencies. Local robustness analysis indicates that measurement error reduces the policymaker’s reaction to model uncertainty, especially at medium and high frequencies.
2 FAMILY LINKAGES AND SOCIAL INTERACTIONS IN HUMAN CAPITAL FORMATION
2.1 Introduction

Human capital is attached to individuals. This feature distinguishes human capital from physical capital but also makes it poor collateral to lenders. Claims against children’s education, knowledge, health, skills, or values are generally not enforceable.\(^1\) Therefore, access to capital markets to finance investments in children is considered to be imperfect in the literature. Family then naturally plays a crucial role in financing human capital accumulation. In his most celebrated work (1964, 1993a, 3rd ed.), Gary Becker develops a (Beckerian) framework modeling family investments in children’s human capital as rational choices of parents who are altruistic toward their children. Differences in human capital and hence in wealth of parents are transmitted to children simply because of distinct financial constraints across families, despite the same level of parental generosity (Becker and Tomes, 1979, 1986). There has been a growing interest in the intergenerational transmission of human capital through family linkages; see Solon (1999) and Black and Devereux (2011) for recent overviews.

To invest in human capital is to invest in human beings. Behavioral responses of children are essential for their acquisition of human capital. They are not passive recipients of parental investments, but active co-investors and/or producers of human capital.\(^2\) The effort that children devote to learning activities largely reflects their attitudes toward education and determines the effectiveness of parental

---

\(^1\)This is because human capital cannot be bought or sold separately from the owner, let alone the immorality of transactions in people (e.g., slavery).

\(^2\)A payment of college tuition, for example, does not automatically translate into college education unless the student wants it. Neither can one deny children’s own desire to change their own lives through education, especially when they are from poor families.
investments. Surprisingly, however, there is not much literature examining the interactions between parents and children in the process of human capital formation. The vast empirical literature on returns to schooling and on-the-job training typically ignores the behavioral reactions of children as rational decision makers; in this aspect, it treats human capital investments no different than investments in physical or financial assets. This essay aims to fill the gap in theory by allowing child effort to enter into the production of human capital so that interesting parent-child interactions arise.

In this extended Beckerian framework, the formation of human capital, jointly determined by parental investment and child effort, is an equilibrium outcome of parent-child interactions. This essay shows that the decentralized equilibrium is Pareto efficient within the family in the sense that neither the parent nor the child can be better off without the other being worse off. However, human capital of the child is a public good within the family; the child enjoys the future income it earns whereas the parent derives utility from it because of altruism.\(^3\) From an instantaneous viewpoint, strategic parent-child interactions lead to potential underproduction of human capital relative to the optimal level chosen by a (fictitious) family planner who maximizes the aggregate family welfare, following the same economic logic as underprovision of public goods by private suppliers. The leading cause in this scenario is that the parent is reluctant to take into full account the benefit of her investment to the family as a whole. The child’s effort choice obeys the same decision rule as in the instantaneous family optimum since his incentive

\(^3\)Notice that the production of human capital is costly to both of them; effort input causes disutility to the child, and the parent has to give up consumption to invest in the child.
schemes are not distorted, given the consistent preferences over education and effort between the parent and himself. This asymmetry in their responsibility for the underproduction result is clearly important for understanding family ties in the transmission of human capital. This essay further argues that a cooperative solution to achieve the family optimum cannot be sustained in the current setting, because there is no mechanism for any transfer from the child to his parent, and neither can the parent borrow from the market against the future income of her child. Therefore, intra-household reallocation by bargaining over the collective outcome of the family optimum, as proposed by Chiappori (1992), is not feasible here across generations. Instead, if one considers a dictatorial parent who can decide an effort level for the child but maximizes her own altruistic utility, parent-child interactions under the liberal parent who does not directly control child behavior will yield the equilibrium output of human capital that coincides with the dictatorial parent’s choice. Although in the same spirit of Becker (1974)’s "Rotten Kid Theorem," this analysis is new to the literature on human capital. It does not suggest that parenting approaches are irrelevant for human capital development; rather, it states that even the dictatorial parent’s decision is inevitably subject to the incentive compatibility constraint of the child.\footnote{Parenting and early childhood intervention have been documented to be rather important, especially when they affect the formation of children’s noncognitive skills that have long-run behavioral consequences, as emphasized by Heckman et al (2006).}

Furthermore, human capital is accumulated within social groups. Interactions with siblings, classmates, neighbors, friends, and other peers heavily shape children’s attitudes toward education and learning activities (especially when driven by
sociological and/or psychological factors such as conformity, imitation, and jealous-
ness). There has been substantial evidence of both sibling effects (e.g., Hauser and
Wong, 1989; Oettinger, 2000) and peer effects (e.g., surveyed by Epple and Romano,
2011) in education. This type of feedback effects within reference groups creates
channels for intragenerational transmission of human capital, which must also
interplay with intergenerational family linkages. Then, some important questions
need to be addressed: How do these two types of forces interplay within families?
And, across families? What are the implications for the dynamics of human capital,
the rise and fall of families, and the evolution of neighborhoods? At the macro
level, a full human capital theory of intergenerational mobility has to account for
the transmission of human capital along both inter- and intra-generational dimen-
sions and the interplay between them; so must a theory of cross-sectional income
distribution and inequality. This essay introduces a social dimension of children’s
behavior into the Beckerian model with endogenous effort choice to pursue this
line of analysis. Two kinds of reference groups are considered successively; one is
siblings within the family, in which the parent exerts direct influences over behavior
of all children, and the other is children from neighboring families, whose actions
are out of the parent’s control unless a choice of neighborhood in which to live
is available. For the convenience of future empirical implementation, this essay
studies a linear quadratic specification of the models throughout, in which the
notion of social multipliers is generalized to conceptualize the role of parent-child
interactions in peer behaviors.

The behavior of children adapts to the interactions with their siblings while
they compete for resources from the parent. In the presence of positive spillovers of effort such as sibling effects on the educational aspirations of children, coordination among siblings will lead to higher levels of human capital output than the non-cooperative decisions for any given parental investments. However, parent-child interactions may offset the spillover effects and give the opposite result. Parent-child interactions can also eliminate potential multiple equilibria with strategic complementarities among siblings. These results largely originate from the structure of intrafamily interactions: a higher layer of interactions between the parent and children is imposed upon the standard inter-sibling game. As asserted by Becker (1993a, p. 21), "No discussion of human capital can omit the influence of families on the knowledge, skills, values, and habits of their children." The analysis with inter-sibling effects developed here enriches the understanding of family linkages in human capital formation.

Human capital transmission also extends beyond families to communities. Within a given neighborhood (which can be a residential community or reference group in a well-defined social space), positive feedback effects among children counteract the pass-through of human capital from parents to children. These impacts on children’s behavior also induce changes in parental investments, which generate a leverage effect that further offsets the influence of families; this is a new insight provided by the model of endogenous parent-child interactions. The intragenerational feedbacks in the neighborhood then potentially accelerate regression towards the mean in the economic status of the families. The theory partially explains why families living in the same community tend to be homogeneous. When the choice of
a group with which to interact (e.g., neighborhood choice) is available, parents will take social interactions into consideration in their neighborhood and investment decisions. It is possible for residential segregation and economic stratification to emerge as equilibrium outcomes so that inequality between communities becomes persistent. This essay characterizes the conditions for the stratified equilibrium in the presence of both intrafamily interactions and neighborhood effects. It shows that there are other channels than strategic complementarities among children to get equilibrium segregation.

As a useful empirical background, Lee and Solon (2009) document that intergenerational income elasticities in the United States are relatively stable around 0.44 over the last two decades of the 20th century. However, many other studies obtain much lower estimates; see Solon (1999) and Black and Devereux (2011) for comprehensive surveys. Recent research has shifted to quantifying the causal relationship between parents and children’s education, for example disentangling the effects of "nurture" from "nature" – predetermined genetics. The evidence is largely mixed, but does often find small causal effects of parental education (Black and Devereux, 2011). On the other hand, research consistently detects positive peer effects in different behaviors of children; see Brock and Durlauf (2001a), Durlauf (2004), and Blume et al (2011a) for identification and evidence of general social interactions and Epple and Romano (2011) for a recent overview on the evidence of peer effects in education specifically.

In theory, since Becker and Tomes (1979, 1986), family linkages for human capital transmission have been further examined in full dynamic settings, for example, by
Aiyagari et al (2002) and incorporated into various macroeconomic models. Cunha et al (2006) go beyond to investigate the family’s role in the whole life cycle process of skill formation and apply this framework to interpret a variety of labor market and behavioral outcomes. Theoretical models of social interactions are usually built on the basis of strategic complementarities among group members (Cooper and John, 1988; Milgrom and Roberts, 1990; Brock and Durlauf, 2001b), which this essay follows. The theory of conformity (e.g., Bernheim, 1994) is an intuitive example of such complementarities, which has been extended and used in many studies (e.g., Ljungqvist and Uhlig, 2000), and is adopted by this essay as a workhorse for generating linear peer influences.

Understanding family and social factors in human development has proven to be of interest in a range of theoretical and applied contexts. However, there is a need to better integrate findings from research on family influences and social interactions in education. The contribution of this essay lies in being one of the first attempts to model at the same time both inter- and intra-generational interactions in the formation of human capital within an innovative but coherent Becker-style framework. Going through the analysis of the interplay between these forces delivers substantial implications for the evolution of families and neighborhoods. This essay is in the spirit similar to, for example, Benabou (1996), Durlauf (1996), and Ioannides (2002), but addresses rather different issues. Assuming neighborhood spillovers in child education through either local public finance of education or sociological effects, Benabou (1996) studies socioeconomic segregation while Durlauf (1996) models persistent income inequality with endogenous residential choices. In
their models, however, there are no direct parental investments in children or the essential parent-child interactions. Ioannides (2002) develops an interesting model of the intergenerational transmission of human capital that reflects both individual choices and neighborhood effects. He introduces local effects into intertemporal family decisions as a technical assumption rather than behavioral consequence of children’s own choices as in this essay. Finally, this essay is mainly a theoretical work, but is written with an intention of being taken to the data in the future.

The rest of this chapter is organized as follows. Section 2.2 extends the Beckerian model by allowing child effort choice in the acquisition of human capital. Section 2.3 considers families with more than one child and investigates the interplay between the inter- and intra-generational transmissions of human capital. Section 2.4 examines the implications of social interactions among children from neighboring families. The case with endogenous neighborhood choice is also explored. Section 2.5 concludes with a discussion of potential future work. Proofs for all propositions are found in the technical Appendix.
2.2 Parent-Child Interactions in the Beckerian Model of Human Capital

Children are recognized as rational forward-looking decision makers in this essay.\(^5\) Even in very early childhood, they are sensitive to parents’ attitudes, words, and actions and react to their attention and care, though often unconsciously.\(^6\) In the acquisition of human capital, some children devote great effort and show much self-discipline, while others seem unable to concentrate on learning activities and tend to drop out of school early. An economist may not be very sympathetic to a pure genetic explanation for such behavioral differences. Children take responsibilities for their own decisions and are usually well aware of this point; differences in their learning behavior are then, to a great extent, the results of choice. As any rational agent, they respond to economic incentives and outside conditions, especially relevant parental decisions toward them. Unlike traditional theories that simply model children as production functions of human capital (mechanically converting physical and time inputs into human capital), this essay treats them as active players in the game of human capital formation with their parents.

The embodiment of human capital in people renders parent-child interactions essential for the development of human capital. Parents can send their children to

\(^5\)From the perspective of a developmental psychologist, the rationality of children should be an increasing function of their age. This essay deals with the benchmark case of full rationality, assuming that the whole pre-adult life collapses into one single period. To fix the idea, it is also helpful to think about teenagers or college students as examples in understanding parent-child interactions.

\(^6\)Such parent-child interactions shape children’s personality traits and noncognitive skills, which, as argued by Heckman et al (2006), have profound consequences through their lifetime.
school or set up study schedules for them, but cannot determine how much effort they make or how much knowledge they want to acquire. Even dictatorial parents do not have perfect control over the actions of their children. Instead, parents have to incorporate children’s strategic responses into their decisions of human capital investments. This section provides a formal treatment of the interactions between parents and children documented by many empirical studies and observed in daily life.

The Benchmark Model

Consider a family of a single parent and her only child. The parent lives one period and the child lives two, childhood and adulthood. The parent is altruistic toward her child and makes a tradeoff between personal consumption and educational investment in the child. The child receives parental investment and exerts effort in study in his childhood, which causes disutility to him. He enjoys the return to his human capital in the adulthood. Suppose that the utility function of the parent is additively separable in her private consumption and in the study effort and adulthood income of her child, given as

$$U(c_p) + a [V(w_c) - K(e_c)], \quad (2.1)$$

where $U(c_p)$ is the parent’s private utility from consumption $c_p$, constrained by her wealth $w_p$, $0 \leq c_p \leq w_p$; $V(w_c)$ is the child’s utility from adulthood income $w_c$, $0 \leq w_c < \infty$; and $K(e_c)$ is the disutility that he suffers from effort input $e_c$. 

0 \leq e_c \leq \hat{e}_c. \quad \text{Further, assume that } U' > 0, U'' < 0, V' > 0, V'' < 0, K' > 0, K'' > 0, \text{ and all satisfy the Inada conditions. Parameter } a, \text{ usually } 0 < a < 1, \text{ is a constant that measures the altruism of the parent. Since the utility of the parent depends on that of the child, formulation (2.1) also implies that the child's utility function is}

\[ V(w_c) - K(e_c). \quad (2.2) \]

It is tempting to interpret \( V(w_c) \) as a value function with state variable \( w_c \) and to think of the model as having an infinite horizon with an overlapping generations structure. This essay does not intend to pursue these recursive dynamics, but focuses on parent-child interactions in the current tractable static framework. One may either interpret \( V(w_c) \) as the present value of utility or as absent of any discounting at all. Whether the child discounts the future or not, parent-child interactions remain unchanged because the parent does the same calculation. There is no loss of generality if the disutility \( K(e_c) \) of effort is normalized as simply \( e_c \) (i.e., measure effort in utils). This essay does not implement this normalization, because, as to be shown soon, it is convenient to specify a linear model with \( K(e_c) \) being a convex function. In addition, utility functions (2.1) and (2.2) have implicitly assumed that the parent and child have consistent preferences over learning effort and future income.\footnote{This essay assumes effort input as a continuous variable. It is also possible to model effort choice as discrete. In this latter case, an interesting equilibrium concept would be a mixed strategy Nash equilibrium, because using mixed strategies to induce parental investment may be able to explain some of children’s behavior in the real world.}

\footnote{One might think that some degree of preference inconsistency between parents and children – for example, the disutility \( K(e_c) \) not entering into the parent’s utility or the child facing extra costs of study effort that the parent ignores – would be necessary for generating a conflict of interest and}
The production of human capital $H_c$ is endogenously determined by two variables: financial investment $0 \leq y_c \leq w_p$ from the parent and effort input $0 \leq e_c \leq \hat{e}_c$ by the child; i.e.,

$$H_c = f(y_c, e_c), \quad (2.3)$$

where the marginal productivity of each factor is positive, $f_y > 0$ and $f_e > 0$, $f(y_c, e_c)$ is concave in $(y_c, e_c)$, $f_{yy} < 0$, $f_{ee} < 0$ and $f_{yy}f_{ee} - f_{ye}^2 > 0$, and the Inada conditions hold. Importantly, parental investment and child effort are technologically complementary to each other, $f_{ye}(y_c, e_c) > 0$ – marginal productivity of each factor rises when the input of the other factor increases. It is natural and also plausible to demand that human capital production should additionally depend on the child’s ability, part of which is inherited from his parent, the parent’s human capital, and other non-negligible family characteristics. Since their impacts are exogenous, these extra factors are embedded in the functional form of $f(y_c, e_c)$ and can be spelled out explicitly if necessary.

For simplicity, suppose that human capital earns a uniform market return $r$ and this is the only income source for the child. Then, his adult wealth $w_c$ is given as

$$w_c = rH_c = rf(y_c, e_c). \quad (2.4)$$

hence strategic interactions between them, but it is not. One might further justify this assumption on the basis of different life experiences between the two generations, arguing that children are more likely to be shortsighted or even irrational and unable to correctly calculate the costs and benefits of study effort from a lifetime perspective. The author is not sympathetic to such arguments. It is hard to imagine that parents do not care about the disutility that their children experience. Neither is there any evidence that children are systematically biased in evaluating the value of education compared to their parents. The consistent preferences (2.1) and (2.2) induce interactive decision-making because the parent does not evaluate the child’s utility the same as he does for himself and the child is not symmetrically altruistic toward his parent.
The model certainly can be generalized to include other sources of income and to allow the parent to directly transfer financial wealth to her child in the form of bequests. But even in this simple environment one can show how the parent and child adjust due to changes in economic incentives (e.g., change in $r$). To complete the description of the model, notice that financial investment $y_c$ has to come from the parent. The child cannot borrow from the market against his future earnings, and neither can the parent. The parent’s budget constraint,

$$c_p + y_c = w_p,$$  \hspace{1cm} (2.5)

then has to be binding.

To solve the model, this essay argues that a Nash equilibrium is the natural solution concept for the fundamental parent-child interactions. As noted above, the child’s behavior toward education cannot be perfectly monitored or controlled. The dictatorial approach of parenting is not an appealing alternative, at least, for the purpose of understanding parent-child interactions. The parent, although altruistic, cares about her own interest and should not be expected to act as a benevolent family planner either. Furthermore, within the current framework, there is an absence of any plausible mechanism for transfers from the child to his parent. The parent cannot receive compensations by issuing claims against human capital of her child. Therefore, intra-household reallocation by bargaining over a collective outcome of the joint production as proposed by Chiappori (1992) and Browning and Chiappori (1998) is not implementable in this intergenerational setting. If there
is no possibility for the parent to share the gain of cooperation, she will deviate to pursue her own interest. A cooperative solution to achieve, for example, the family planner’s optimum then cannot be sustained.

Taking child effort $e_c$ as given, the parent chooses consumption $c_p$ and investment $y_c$ to maximize her utility (2.1) subject to constraints (2.4) and (2.5). Under the assumptions of the model, she has an interior solution determined by

$$rf_y(y_c, e_c) = \frac{U'(c_p)}{aV'(w_c)}, \quad \text{for any } e_c \in [0, \hat{e}_c].$$

Condition (2.6) requires that the gross return to parental investment $R_1(y_c, e_c) \equiv rf_y(y_c, e_c)$ equals the marginal rate of substitution. This is the standard principle for utility maximization extended to the case of an altruistic agent. In light of this condition, the parent will invest more in human capital when (i) she is more altruistic ($a$ is higher), or (ii) she is richer ($w_p$ is higher). However, caution is needed in drawing a conclusion when the gross return $R_1$ changes, which depends on market return $r$, child effort $e_c$, and the production technology. A change in $R_1$ induces a substitution effect as well as an income effect (notice that $R_1$ also affects $V'(w_c)$ in equation (2.6)), leaving the total effect undecided. On the other hand, the child’s choice of effort $e_c$ to maximize utility (2.2) is optimal when the gross return to effort input $R_2(y_c, e_c) \equiv rf_e(y_c, e_c)$ equals his marginal rate of substitution.

---

9Wealthier parents are very likely to have received more education, so case (ii) suggests the possibility of intergenerational transmission of human capital.
between effort and income,

\[ rf_{c}(y_{c}, e_{c}) = \frac{K'(e_{c})}{V'(w_{c})}, \text{ for any } y_{c} \in [0, w_{p}]. \]  \hspace{1cm} (2.7)

The effects of an increase in market return \( r \) or parental investment \( y_{c} \) are again twofold.

A profile of investment and effort choices, \((y_{c}^{N}, e_{c}^{N})\), is a pure strategy Nash equilibrium of the benchmark model outlined above if it satisfies both conditions (2.6) and (2.7). And there exists such a Nash equilibrium; appendix A.1 establishes the existence of the equilibrium, but it does not exclude the possibility of multiple equilibria. For example, it is entirely possible for both the parent and child being aggressive toward human capital accumulation to be one equilibrium and both shirking to be another equilibrium, under certain conditions.

To understand equilibrium behavior, it is helpful to examine the reaction functions of the parent and child more closely. Notice that the parent’s reaction to child effort, \( y_{c} = \phi_{1}(e_{c}) \), is an implicit function contained in equation (2.6) and, at the margin,

\[ \phi_{1}'(e_{c}) = -\frac{ar(V'f_{ye} + rf_{y}f_{e})}{U'' + ar(V'f_{yy} + rf_{y}f_{e})}, \]  \hspace{1cm} (2.8)

where \( U'' + ar(V'f_{yy} + rf_{y}f_{e}) < 0 \), \( arV'f_{ye} > 0 \), and \( ar^2V''f_{y}f_{e} < 0 \). When the child is working harder, the marginal return to parental investment increases due to the technological complementarity. But it also allows the parent to consume more while still maintaining a similar level of human capital as before, which undermines her incentive to invest. The sign of the parent’s marginal reaction
(2.8) is then determined by the dominance of the price/substitution effect or the income/crowding-out effect. The reaction function of the child, $e_c = \phi_2(y_c)$, is implied by condition (2.7) and has

$$\phi_2'(y_c) = \frac{r (V'f_{ey} + r V''f_{ey}f_{y})}{K'' - r (V'f_{ee} + r V''f_{e}^2)},$$

with $K'' - r (V'f_{ee} + r V''f_{e}^2) > 0$, $r V'f_{ey} > 0$, and $r^2 V''f_{e}f_{y} < 0$. The intuition again follows: parental investment increases the productivity of child effort while directly substituting for the effort itself. One can conclude from equations (2.8) and (2.9) that the assumptions imposed so far cannot decide the number of fixed points of the function $\phi \equiv [\phi_1, \phi_2]'$ in set $\Pi \equiv [0, w_p] \times [0, \hat{e}_c]$ that define Nash equilibria. It is also interesting that the marginal reaction functions (2.8) and (2.9) have the same sign that is determined by the same condition, as summarized in the following proposition.

**Proposition 2.1.** If technological complementarity between parental investment $y_c$ and child effort $e_c$ is strong enough such that

$$\frac{f_{ye}}{rf_{y}f_{e}} > -\frac{V''}{V'}, \quad \text{for any } (y_c, e_c) \in \Pi,$$  

then the parent and child react positively to each other, $\phi_1'(e_c) > 0$ and $\phi_2'(y_c) > 0$.

Proposition 2.1 describes the property of parent-child interactions under strong technological complementarity. Inequality (2.10) requires that the objective complementary effect in productivity normalized by the first order effects exceeds the
subjective value of the income effect measured by the curvature of $V(w_c)$. It guarantees that the substitution effect dominates the income effect whenever the other’s behavior changes. The condition can hold, for example, when market return $r$ is low so that the income effect is always weak. One reason why this proposition is important is because the sign of the marginal reaction functions helps predict the shift of the Nash equilibrium when some exogenous shock hits the family. For example, suppose instead that condition (2.10) fails to hold and so $\phi'_1(e_c) \leq 0$ and $\phi'_2(y_c) \leq 0$ around Nash equilibrium $(y^N_c, e^N_c)$. Then, when an exogenous change in economic environment increases parental investment (e.g., a rise in household income shifts the parent’s reaction curve toward more investment), the child will reduce his effort in the equilibrium. According to the present model, however, this is not the child’s behavioral problem but rather his rational response to economic incentives. Symmetrically, an exogenous shift toward higher child effort results in less resources invested in child education from the parent when $\phi'_1(e_c) \leq 0$ and $\phi'_2(y_c) \leq 0$. This type of symmetry in comparative statics predicted by the theory is new to the literature.

Parents exert considerable influence on their children’s development. In this extended Beckerian framework, the parent’s human capital is transmitted to her child through three distinct channels: (i) parental human capital is an important determinant of the production of child human capital (reflected in the functional form of (2.3) and captures impacts such as role model effects); (ii) its financial returns are directly invested in the child for his education and learning activities; and (iii) parental investment also affects the child’s own effort choice and may even
amplify the direct impacts of investment.\textsuperscript{10} This essay complements the previous literature by putting forward mechanism (iii). The intergenerational transmission of human capital through endogenous family ties involves a rather complicated nonlinear process, which needs to be considered in any careful empirical study of intergenerational mobility.

What are the welfare implications of parent-child interactions? Is the Nash equilibrium Pareto efficient? May the strategic actions fail to reach a state that is optimal for the whole family, at least from an instantaneous viewpoint? How does the decentralized equilibrium compare to the decision of a dictatorial parent who can control child behavior? To address these issues, first, notice that Pareto efficiency within the family simply means that neither the parent nor the child can be better off without the other being worse off. Second, a family planner, if one exists, will choose parental investment $y_c$ and child effort $e_c$ jointly to maximize the instantaneous aggregate welfare function

$$W(y_c, e_c) = U(c_p) + (1 + a) [V(w_c) - K(e_c)],$$ (2.11)

\textsuperscript{10}Parental human capital may also have an effect on child preferences. For example, it is possible for a child of better educated parents to have higher disutility of effort due to the better opportunity of consumption or lower disutility because of the better guidance from the parents. This essay abstract from these considerations assuming homogeneous preferences. Of course, the child also inherits genes from his parent, which largely determine his abilities, health, and attitudes towards education. This is also implicitly included in (2.3), but the current essay does not explicitly model gene/ability inheritance to keep the model simple.
subject to constraints (2.3)-(2.5). Assuming an interior solution results in

\[ rf_y(y_c, e_c) = \frac{U'(c_p)}{(1 + a)V'(w_c)}, \tag{2.12} \]

\[ rf_e(y_c, e_c) = \frac{K'(e_c)}{V'(w_c)}. \tag{2.13} \]

If this family utility \( W(y_c, e_c) \) is concave in \((y_c, e_c)\), then the second order conditions hold automatically and the solution is unique.\(^{11}\) For convenience, denote the relations implied by (2.12) and (2.13) as \( y_c = \chi_1(e_c) \) and \( e_c = \chi_2(y_c) \), respectively. A straightforward representation of conditions (2.12) and (2.13) reveals that the marginal rates of transformation equal the marginal rates of substitution for the family as a whole at the *instantaneous family optimum*. Finally, the dictatorial parent also chooses investment \( y_c \) and effort \( e_c \) together, but to maximize her own utility (2.1) under constraints (2.3)-(2.5). The first order conditions for her decision are consistent with equations (2.6) and (2.7). Then, the following proposition holds.

**Proposition 2.2.** Assume strong technological complementarity as (2.10) and strict concavity of aggregate family utility \( W(y_c, e_c) \) in \((y_c, e_c)\). Then, (i) the Nash equilibrium \((y_c^N, e_c^N)\) is unique and Pareto efficient; (ii) the instantaneous family optimum \((y_c^*, e_c^*)\) is also unique, and \((y_c^N, e_c^N) < (y_c^*, e_c^*)\), i.e., the Nash equilibrium \((y_c^N, e_c^N)\) is not optimal for the family from the instantaneous viewpoint; and (iii) the optimal choice of the dictatorial parent \((y_c^D, e_c^D)\) is unique as well, and it coincides with the Nash equilibrium \((y_c^N, e_c^N)\).

To argue the Pareto efficiency of the Nash equilibrium, it is important to notice

---

\(^{11}\)When the family utility \( W \) is written as a function of \((y_c, e_c)\), this essay assumes that all constraints have already been substituted in and hence only endogenous variables matter for \( W \).
the proportional altruism in the benchmark model. Under this formulation, the Nash choice function of child effort is still valid for Pareto allocation. Then, the Pareto choice of parental investment that guarantees at least the same utility for the child as in the Nash equilibrium and maximizes parental utility simply coincides with the Nash choice \( y_c^N \). Specifically, the Pareto allocation is a choice of \((y_c, e_c)\) that maximizes (2.1) subject to

\[
V(w_c) - K(e_c) \geq V(w_c^N) - K(e_c^N),
\]

(2.14)

and constraints (2.3)-(2.5). Given the form of the extra constraint (2.14), effort choice \( e_c = \phi_2(y_c) \) determined by (2.7) in the Nash equilibrium is still a part of the solution to this Pareto problem. For optimization, constraint (2.14) needs to be binding. Under the technical assumptions on preferences and technologies, \((y_c^N, \phi_2(y_c^N))\) solves (2.14) with equality and the whole problem. Therefore, the Nash equilibrium \((y_c^N, e_c^N)\) is Pareto efficient.

Part (ii) of Proposition 2.2 indicates that parent-child interactions may violate the instantaneous optimality for the family. In the Nash equilibrium, conditions (2.6) and (2.7) imply that

\[
\frac{f_y(y_c^N, e_c^N)}{f_e(y_c^N, e_c^N)} = \frac{U'(c_p^N)}{aK'(e_c^N)} > \frac{U'(c_p^N)}{(1 + a)K'(e_c^N)}.
\]

The marginal rate of transformation between \( y_c \) and \( e_c \) equals the parent’s private marginal rate of substitution, which is greater than the family’s. The parent cares about the disutility that her child suffers because of altruism, but not the full cost
of child effort to the family. From her perspective, parental investment creates a positive externality that benefits the child. Decentralized decision making in the family fails to internalize this externality associated with human capital formation. Proposition 2.2 further shows that there is potentially an underproduction of human capital in the Nash equilibrium relative to the family optimum. Given the public good property of human capital in the family, this is simply the classic result that voluntary provision of public goods is inadequate. Although it is not surprising to have the standard economic reasoning extended to this intrafamily environment, as far as the author knows, this view on human capital formation is novel in the literature. Interestingly, the parent’s strategic behavior is the leading cause of the underproduction result. The child behaves exactly the same in the decentralized and centralized settings; his reaction function (2.7) coincides with the relation (2.13) because his incentive schemes are not distorted under the consistent preferences between the parent and child. However, the parent’s behaviors are systematically different under the two regimes. The child simply takes her decisions into account in a consistent way.

Admittedly, Part (ii) of Proposition 2.2 takes a very specific perspective on family welfare by using criterion (2.11). Function (2.11) is a simple cross-sectional aggregation of (2.1) and (2.2) and stands for an instantaneous viewpoint.\footnote{The analysis can be easily generalized to the case of weighted family utility.} For example, it ignores the parent’s disutility of effort that she exerted when she was a child. In a fully dynamic model with overlapping generations, since individuals born at different times attain different utility levels, it is not completely clear how
to evaluate social welfare (Diamond, 1965). The current approach may seem less appealing in a fully dynamic environment, but it is entirely appropriate for the static framework constructed here. And there is no doubt that the instantaneous view is indeed an important angle for thinking about parent-child interactions.

Part (iii) of Proposition 2.2 claims that the production of human capital is identical under either the liberal parent or the dictatorial parent when responding to the same economic conditions. In principle, this is similar to Becker (1974)'s "Rotten Kid Theorem," which states that if a family has a caring household head who gives money to all other household members, then each member, no matter how selfish, will maximize his or her own utility by taking actions that lead to maximization of total family income. Thus, all family members will act harmoniously in the family interest. Proposition 2.2 uncovers the same insight that parent-child interactions lead to the equal outcome as the parental dictation in human capital development, which is a rather distinct context from Becker (1974). In addition, the result of part (iii) does not imply that parenting approaches are irrelevant for human capital formation, but insists that even the dictatorial parent’s decision is subject to the child’s incentive conditions.

In sum, Proposition 2.2 provides some new thoughts on family management in the formation of human capital. Family structures are quite different in the real world: some families are fairly decentralized, while some have more features of dictation. No matter how parents manage the behavior of their children, they have to take children’s incentive constraints into consideration in their investment decisions, which is the essential point of parent-child interactions, and the outcome
is Pareto efficient within the family. As noted above, the suboptimality of parent-child interactions from the family’s viewpoint derives from the generic constraints in financing human capital investments imposed by the nature of human capital. The present analyses go beyond the standard method that treats a household as a single, unified entity in the characterization of family decisions.

As a technical matter, it is essential in Proposition 2.2 to assume that the family aggregate utility $W(y_c, e_c)$ is concave in $(y_c, e_c)$. This assumption is not only necessary for the uniqueness of solutions but also for the underproduction result. To see the idea, notice that concavity requires the determinant of the second-order derivative matrix to be positive, which, as verified Appendix A.3, gives the following condition:

$$\chi'_1(e_c) = -\frac{(1 + a)r (V'f_{ye} + rV''f_{f}f_{e})}{U'' + (1 + a)r (V'f_{yy} + rV''f_{f}^2)} < \frac{K'' - r (V'f_{ee} + rV''f_{e}^2)}{r (V'f_{ye} + rV''f_{f}f_{y})} = \frac{1}{\chi'_2(y_c)},$$

where the complementarity condition (2.10) imposes that $V'f_{ye} + rV''f_{f}f_{e} > 0$. As illustrated in Figure 2.1, this condition means that the inverse function $y_c = \chi^{-1}_2(e_c)$ is always steeper than function $y_c = \chi_1(e_c)$. Therefore, they can cross only once and yield the unique family optimum. The parent’s reaction curve $y_c = \phi_1(e_c)$ in the decentralized setting always lies below $y_c = \chi_1(e_c)$, and it is increasing in effort $e_c$ but more slowly than $y_c = \chi_1(e_c)$, i.e., $\phi'_1(e_c) \leq \chi'_1(e_c)$. The child’s reaction $e_c = \phi_2(y_c)$ is the same as relation $e_c = \chi_2(y_c)$. Then, the two decentralized reaction curves can only intersect once as well, giving the unique Nash equilibrium, and the intersection $(y_c^N, e_c^N)$ lies to the left of and below the instantaneous family
Figure 2.1: Nash Equilibrium and Family Optimum under Complementarity Condition (2.10)

Figure 2.2: Nash Equilibrium and Family Optimum under Complementarity Condition (2.15)
optimum \((y_c^*, e_c^*)\), indicating the underproduction of human capital.

Finally, the result of Proposition 2.2 on decentralized decisions and family welfare relies on technical assumptions, especially condition (2.10). What if the opposite holds? When

\[
\frac{f_{ye}}{ff_{ye}} < -\frac{V''}{V'''} \quad \text{for any } (y_c, e_c) \in \Pi,
\]

(2.15)

the relation \(e_c = \chi_2(y_c)\) for the family optimum still agrees with the reaction function of the child \(e_c = \phi_2(y_c)\), and its inverse curve \(y_c = \chi_2^{-1}(e_c)\) is steeper than the parent’s behavioral curves \(y_c = \chi_1(e_c)\) and \(y_c = \phi_1(e_c)\) in both cases if one maintains the concavity of \(W(y_c, e_c)\). It immediately follows again that there is a unique instantaneous family optimum and Nash equilibrium. However, condition (2.15) renders all behavioral functions downward sloping. As displayed in Figure 2.2, parental investment is higher while child effort is lower in the family optimum than in the Nash equilibrium. It is then uncertain which solution results in a higher level of human capital.

A Linear Quadratic Specification

This section provides an analysis based on a linear quadratic specification of preferences and technologies in order to have tractable solutions and use them to make predications. This type of linear quadratic model has been widely studied in the economic literature (cf., Hansen and Sargent, 2005). For empirical purposes, the analysis also suggests a potential strategy for taking the model to the data in follow-
up research. In particular, since the seminal work of Manski (1993), linear-in-means models of social interactions have been intensively explored, concerning their identification issues and applications in empirical work. Blume et al (2011a, 2011b) show that linear behavioral equations involved in this body of literature can be derived from micro-founded linear quadratic models. Given that much of the interest lies in behavioral functions (especially when a social dimension of child behavior is introduced later), this section follows their approach to specify the benchmark model outlined above. Despite analytical results, some interesting properties – for example, multiple equilibria – may be lost in the linear specification.

Specifically, let the parent’s private utility from consumption be

$$U(c_p) = \theta_1 c_p - \frac{\theta_2 c_p^2}{2}$$

(2.16)

where $\theta_1 > 0$, $\theta_2 > 0$, and $U'(w_p) = \theta_1 - \theta_2 w_p > 0$. The production function of human capital $H_c$ is linear in parental investment $y_c$,

$$H_c = A_c y_c,$$

(2.17)

but productivity $A_c$ of investment depends on child effort $e_c$, ability $a_c > 0$, parental human capital $H_p$, and other important characteristics $x_h$ of the child,

$$A_c = \kappa e_c + \rho a_c + \sigma H_p + \sum_h \beta_h x_h.$$

(2.18)

Reasonable values of parameters $\kappa$, $\rho$, and $\sigma$ are usually nonnegative. Equation (2.18)
then allows one to control for relevant variables other than parental investment 
and child effort in an empirical study. On the other hand, suppose that the child’s 
utility from adulthood income $w_c$ is

$$V(w_c) = \eta_1 w_c - \frac{\eta_2}{2} \left( \frac{w_c}{r A_c} \right)^2, \quad (2.19)$$

with $\eta_1 > 0$ and $\eta_2 > 0$. Since his human capital earns income $w_c = r H_c$, $r > 0$, 
expression $\frac{w_c}{r A_c}$ in (2.19) simply reduces to $y_c$, the level of parental investment. 
Letting $V(w_c)$ be quadratic in $y_c$ seems restrictive, but it is necessary for generating 
a linear behavioral function for the child. Still, this quadratic term can be interpreted 
as the cost of human capital investment and hence rationalized to be independent 
of any characteristics of the child. The child’s disutility of effort $e_c$ is

$$K(e_c) = \alpha e_c + \frac{1}{2} e_c^2, \quad (2.20)$$

with $\alpha > 0$. Finally, the budget constraint (2.5) is still effective.

Under this simple specification, the parent’s response function $y_c = \phi_1(e_c)$ is 
linear in child effort $e_c$,

$$y_c = -\theta_1 + \theta_2 w_p + a \eta_1 r A_c, \quad (2.21)$$

The coefficient on effort $e_c$ depends on the degree of altruism $a$, the marginal utility 
of income for the child $\eta_1$, the market return to human capital $r$, and the techno-
logical complementarity between investment and effort $\kappa$. She also proportionally
increases her investment in child education when the marginal utility of private consumption is lower and the wealth endowment is higher. Similarly, the child’s decision rule \( e_c = \phi_2(y_c) \) is also linear in \( y_c \),

\[
e_c = -\alpha + \eta_1 r \kappa y_c. \tag{2.22}
\]

His effort input is lower when the marginal disutility of effort is higher and parental investment is lower. The latter effect depends on the value of future income, the market return, and the production technology.

Assume that there exists an interior Nash equilibrium \( (y_c^N, e_c^N) \in \Pi \). Then it must be unique in this linear environment and can be solved from equations (2.21) and (2.22) as

\[
y_c^N = \frac{B_c - \sigma \alpha}{\eta_1 r \kappa (1 - \sigma)}, \quad \text{and} \quad e_c^N = \frac{B_c - \alpha}{1 - \sigma}, \tag{2.23}
\]

where \( \sigma \equiv \frac{\alpha (\eta_1 r \kappa)^2}{\theta_2 + \alpha \eta_2} \) and \( B_c \equiv \frac{\eta_1 r \kappa}{\theta_2 + \alpha \eta_2} \left( (\theta_2 w_p - \theta_1) + \alpha \eta_1 r (\rho a_c + \sigma H_p + \sum_h \beta_h x_h) \right) \).

Intuitively, \( B_c \) can be interpreted as a comprehensive but still linear aggregation of the exogenous economic factors of the parent (the first part in the brackets) and the child (the second part). Consider the interesting case that \( 0 < \sigma < 1 \) and hence \( B_c > \alpha \). First, writing the equilibrium parental investment \( y_c^N \) in terms of the exogenous factors of the family as (2.23), one can immediately see the endogeneity problem in the traditional empirical research that takes the choice equation of \( y_c \) with fixed \( e_c \) as (2.21) directly to the data. Second, the direct effect of any change in exogenous variables on equilibrium investment or effort is amplified by the multiplier \( \frac{1}{1-\sigma} \).

This leveraging effect of parent-child interactions has important implications for the
rise and fall of families. The converging force toward homogeneity in the economic status across families driven, for example, by peer interactions among children can be further amplified by parent-child interactions, which will then accelerate the converging process; a detailed model will be presented in Section 2.4.

Given parental investment and child effort in (2.23), the output of human capital $H_N^c$ in the Nash equilibrium is

$$H_N^c = \frac{(B_c - \sigma \alpha)^2}{\eta_1 r \sigma (1 - \sigma)^2} + \frac{(\theta_1 - \theta_2 w_p) (B_c - \sigma \alpha)}{\alpha \kappa (1 - \sigma)(\eta_1 r)^2}.$$  \hspace{1cm} (2.24)

If parental wealth $w_p$ is endowed in the form of human capital, i.e., $w_p \equiv r H_p$ in $B_c$, then the intergenerational transmission of human capital is nonlinear even in this simplest linear setting. $B_c$ is raised to the power of two when entered into expression (2.24), and so is the multiplier $\frac{1}{1 - \sigma}$. The main message of this expression is that any serious empirical implementation of the human capital theory of intergenerational mobility has to model parent-child interactions carefully. A relevant empirical question to ask here is how to measure the leverage multiplier of parent-child interactions in the data. Even if one can estimate equation (2.24), an identification strategy is still necessary to uncover the multiplier $\frac{1}{1 - \sigma}$. This topic is interesting, but beyond the scope of this chapter.

### 2.3 Parental Investments and Sibling Interactions

Decision making in group contexts necessarily reflects the influence of groups. Beside the same contextual factors that all group members are exposed to, more
importantly, there are endogenous effects in their behavior due to conformity, group identification, imitation, learning spillovers, or even jealousness. Siblings form a natural reference group within the family for children. Research has found substantial correlation between siblings not only in education (e.g., Oettinger, 2000) but also in risky behavior (e.g., Altonji et al., 2010), which affects human capital formation indirectly, and has been able to partly establish the causal relationship (e.g., Hauser and Wong, 1989). Endogenous sibling effects create horizontal family linkages for the transmission of human capital, in addition to the vertical interactions between parents and children. This section develops an analysis of both inter- and intra-generational family linkages in the process of human capital formation, which can also be interpreted as a new model of social interactions with endogenous contextual effects; to a great extent, the contextual effects among siblings depend on the decisions of their parents and hence should not be treated as exogenous.

**The Model with Sibling Effects**

There must be multiple children in a family to introduce sibling interactions. Consider the same setup as Section 2.2 except that the parent now has \( n \) children, \( n > 1 \). Following Becker (1993a), let the parent be equally altruistic toward each child.\(^{13}\)

\[^{13}\text{It may be also of interest to consider that the parent has a minmax preference over her children,}
\]

\[ U(c_p) + \min_i \left[ V(w_{c_i}^i) + S(e_{c_i}^i, \bar{e}_c - e_{c_i}^i) - K(e_{c_i}^i) \right]. \]

That is, her altruistic utility component is solely determined by the child of the lowest utility. Under this preference, the parent tries to maintain a sort of equality among her children.
Then, her problem can be summarized as

$$\max_{\{c_p, y_c^i\}} \left\{ U(c_p) + a \sum_{i=1}^{n} \left[ V(w^i_c) + S(e^i_c, \bar{e}^{-i}_c) - K(e^i_c) \right] \right\} \quad (2.25)$$

subject to

$$w^i_c = rH^i_c = rf^i(y^i_c, e^i_c) \quad (2.26)$$

$$c_p + \sum_{i=1}^{n} y^i_c = w_p, \quad (2.27)$$

taking the effort choices of her children $e^i_c, i = 1, \ldots, n$, as given. Clearly, the new utility component $S(e^i_c, \bar{e}^{-i}_c)$ to be explained later is irrelevant for her decision. The production function of human capital $f^i(\cdot, \cdot)$ is indexed with superscript $i$, so it can include heterogeneous abilities and relevant individual characteristics of children. This is actually the only source of heterogeneity among children in this section.\(^\text{14}\)

The first order conditions for the optimization problem (2.25)-(2.27) are

$$rf^i_y(y^i_c, e^i_c) = \frac{U''(c_p)}{aV'(w^i_c)}, \quad \text{for any } e^i_c \in [0, \bar{e}_c], \text{ and } i = 1, \ldots, n. \quad (2.28)$$

The gross return $R^i_1(y^i_c, e^i_c) \equiv rf^i_y(y^i_c, e^i_c)$ to investment in child $i$ must equal the parent’s marginal rate of substitution between her own consumption and child $i$’s future income. This does not mean that investment in each child earns the same return,\(^\text{15}\) or the parent should make the same investment in each child.

\(^{14}\)They may have different preferences, which causes their differential effort. But the nature of interactions remains the same as here. Although it is also possible for parents to manipulate the preferences of their children as studied in Becker \textit{et al} (2011), this chapter does not intend to pursue this generalization.

\(^{15}\)There is heterogeneity among return functions $R^i_1(y^i_c, e^i_c)$, and $R^i_1(y^i_c, e^i_c)$ also enters $V'(w^i_c)$ in equation (2.28).
Should the parent invest more in children who are working harder, better motivated, or have more talent? The answer is "yes" for higher returns in the classic Beckerian model absent of children’s active roles, and "yes, if the technological complementarity is strong" in the single-child model of Section 2.2. In this multi-child setting, however, condition (2.10) is not sufficient for child i to receive more parental investment when he pledges more effort. To show this, differentiate condition (2.28) with respect to $e_i^c$:

$$
\frac{\partial y_i^c}{\partial e_i^c} = -\frac{\alpha r \left(V'f_{ye}^i + rV''f_y^if_e^i\right) + U'' \sum_{j \neq i} \frac{\partial y_j^c}{\partial e_i^c}}{U'' + \alpha r \left(V'f_{yy}^i + rV''f_y^i f_e^i\right)^2}.
$$

Then it is clear that the effects of an increase in child effort $e_i^c$ are threefold: (i) it makes investment in child i more productive due to the technological complementarity $y_{ye}^i > 0$, and hence raises the parent’s incentive to invest in him; (ii) it has an income effect $V'' < 0$ since increased effort directly substitutes for investment, which undermines the incentive; and (iii) there is potential substitution across children because the parent may compensate or punish his siblings for lower effort inputs (ambiguous sign of $\sum_{j \neq i} \frac{\partial y_j^c}{\partial e_i^c}$), which gives an uncertain effect. Effects (i) and (ii) are the same as in the single-child family, and their sum is positive when condition (2.10) holds for i. Effect (iii) is new; for example, when it dominates the other effects,

$$
\sum_{j \neq i} \frac{\partial y_j^c}{\partial e_i^c} > -\frac{\alpha r}{U''} \left(V'f_{ye}^i + rV''f_y^i f_e^i\right),
$$

the parent will respond negatively to the child’s effort increase, $\frac{\partial y_i^c}{\partial e_i^c} < 0$. Children will certainly take the parent’s behavior into consideration in any full equilibrium.
The interesting point is that in the framework of endogenous parent-child interactions, siblings are forced to interact with each other, though indirectly through their parent’s decision, even when there is an absence of any direct connections (e.g., $S(e^i_c, \bar{e}^{-i}_c)$ in (2.25)) between their own choices. The driving force is to compete for resources from the parent. As pointed out by the above example, Proposition 2.1 does not always fit multi-child families. This complication may seem unpleasant, but it allows for the analysis of multidimensional intrafamily interactions in the real world, which is essential for a full understanding of human capital formation within families.

To model endogenous interactions among siblings, this essay, following Brock and Durlauf (2001b)’s approach, introduces a group component of utility $S(e^i_c, \bar{e}^{-i}_c)$ associated with effort choices into the utility function of children. Explicitly, write the children’s utility function as

$$V(w^i_c) + S(e^i_c, \bar{e}^{-i}_c) - K(e^i_c), \quad (2.29)$$

where $\bar{e}^{-i}_c \equiv \frac{1}{n-1} \sum_{j \neq i} e^j_c$ is the exclusive average of siblings’ effort inputs. This group utility $S(e^i_c, \bar{e}^{-i}_c)$ can be interpreted as a conformity effect, a formulation of sibling pressure, or simply learning spillovers among children, which can be quite strong within the family. First of all, the expression (2.29) has assumed that only the average choice of siblings matters for the group utility. One can imagine possible asymmetric influences between siblings; for example, spillovers of effort

---

16If the parent’s utility function is such that optimal consumption is constant, then condition (2.10) is sufficient for the parent’s positive reaction to child i’s effort input, because now $\sum_{j \neq i} \frac{\partial y}{\partial e^j_c} = -\frac{\partial y}{\partial e^i_c}$. 
may be stronger in one direction than the other. Some empirical studies find that
gender and birth order can also affect the magnitude of sibling influence. Interested
in reciprocal influence, this essay abstracts away from these details. Hauser and
Wong (1986), for example, find evidence of reciprocal influence of brothers’ levels of
educational attainment, net of common effects of family background and the effect
of each brother’s mental ability on his schooling, and show that a model of equal
reciprocal effects fits the data almost as well as a model assuming unconstrained
effects. Second, the group utility component is additive to the child’s private utility
in (2.29). In this regard, it is possible to make alternative assumption; for example,
one can include group effects in the production of human capital or embed the
group utility in the objective function nonlinearly. This essay chooses formulation
(2.29) to keep the model simple, but also because it is standard in the relevant
literature. Third, this essay only deals with the interesting case in which the

\[ \frac{\partial^2 S(w_i c, \bar{w} - ic)}{\partial w_{ic} \partial \bar{w}_{ic}} > 0 \]

This type of "keeping up with the Joneses" utility function has been used by, for example, Ljungqvist and Uhlig (2000), and can be interpreted
as a formalization of jealousness among agents. Under this specification, the children’s decisions
are given by

\[ rf_i^e(y_{ic}, e_{ic}) = \frac{K'(e_{ic})}{V'(w_{ic}) + S_1(w_{ic}, \bar{w}_{ic})}, \quad \text{for any } y_{ic} \in [0, w_p], \text{ and } i = 1, \ldots, n, \]

instead of (2.31). In either case, the basic point is that the effort choice of a child depends on parental
investments as well as the choices of his siblings. The two approaches are the same in essence, even
though they represent different modeling styles. This essay is based on the specification of (2.29).
group utility exhibits strategic complementarity,\(^{19}\) i.e.,

\[
\frac{\partial^2 S(e^i_c, \bar{e}^{-i}_c)}{\partial e^i_c \partial \bar{e}^{-i}_c} > 0.
\]  

(2.30)

Hence, the marginal group utility of child i’s own effort input increases when the average effort of his siblings is higher. Further, assume that \(S(e^i_c, \bar{e}^{-i}_c)\) is concave in \((e^i_c, \bar{e}^{-i}_c)\), with \(S_{11} < 0, S_{22} < 0, S_{11}S_{22} - S_{12}^2 > 0\), and that the Inada conditions hold. In the literature of social interactions, it is common to assume a constant cross-partial \(J\) which allows one to measure the degree of dependence across group members with a single parameter; see the linear quadratic model in Section 2.3 for an example. At this point, there are no restrictions on the signs of derivatives \(S_1(e^i_c, \bar{e}^{-i}_c)\) and \(S_2(e^i_c, \bar{e}^{-i}_c)\). However, if \(S_2(e^i_c, \bar{e}^{-i}_c) > 0\), then there are positive spillover effects (which are distinguished from strategic complementarities that act on marginal utilities).

Given their preferences as in (2.29), children’s decisions are then implicitly characterized by the first order conditions

\[
rf^i_c(y^i_c, e^i_c) = \frac{K'(e^i_c) - S_1(e^i_c, \bar{e}^{-i}_c)}{V'(w^i_c)}, \quad \text{for any } y^i_c \in [0, w_p], \text{ and } i = 1, \ldots, n. \quad (2.31)
\]

Whether a child’s effort is higher or lower than in the absence of siblings influences \((i.e., \text{determined by condition (2.7)}\), holding all the other factors constant, should depend on the sign of \(S_1(e^i_c, \bar{e}^{-i}_c)\), or, more broadly, the specification of group utility.

---

\(^{19}\)This type of complementarity imposes restrictions on preferences and leads to direct strategic interactions in decision making. It is distinct from technological complementarity assumed for the production function of human capital.
S(e^i_c, \bar{e}^{-i}_c).

In this model of multiple children, a profile of investment and effort choices, \((y^{i,N}_c, e^{i,N}_c)_{i=1}^n\), is a pure strategy Nash Equilibrium if it satisfies both conditions (2.28) and (2.31), and there exists such a Nash Equilibrium. Notice that a strategy profile lives in space \(\Pi^n \equiv ([0, w_p] \times [0, \hat{e}_c])^n\). The parent’s investment \(y^i_c\) in child \(i\) depends on the child’s effort \(e^i_c\) and also her investments in other children, \(y^j_c, j \neq i\), in (2.28); denote it as \(y^i_c = \phi^1_i(e^i_c, y^{-i}_c)\). The child reacts to parental investment \(y^i_c\) and effort inputs of his siblings \(e^j_c, j \neq i\), implicitly in equation (2.31); denote his decision function as \(e^i_c = \phi^2_i(y^i_c, e^{-i}_c)\) or specifically, \(\phi^2_i(y^i_c, \bar{e}^{-i}_c)\). Let \(\phi \equiv [\phi^1_1, ..., \phi^1_n, \phi^2_1, ..., \phi^2_n]^t\). Then one can establish a fixed point of function \(\phi: \Pi^n \to \Pi^n\) in the space \(\Pi^n\); see Appendix A.4 for details. Given the assumptions on preferences and technologies, function \(\phi\) is continuous and hence guarantees a fixed point that defines a Nash equilibrium.

As is evident from function \(\phi\), children interact with each other and with their parent in a rather complicated way. An additional question naturally arises here: how do the noncooperative interactions among children, compared with their joint decision, shape parent-child interactions? To address this issue, parent-child interactions have to be present; hence it is not a full cooperative equilibrium even with cooperation among children. To simplify the analysis, the rest of this section assumes away heterogeneity in children so that the production functions \(f^i(\cdot, \cdot)\) are the same across \(i\). It is without loss of generality in this scenario to focus on symmetric equilibria.
In the noncooperative equilibrium, the children’s choice \( e_c \) as a function of parental investment \( y_c \), \( e_c = \varphi^N_2(y_c) \), is given by the symmetric version of (2.31),

\[
rf_e(y_c, e_c) = \frac{K'(e_c) - S_1(e_c, e_c)}{V'(w_c)}.
\] (2.32)

In contrast, the cooperative decision making among children also takes the average effort \( \bar{e}_c^{-i} \) in \( S(e_{c}^i, \bar{e}_c^{-i}) \) as a choice variable in the maximization of the representative utility (2.29). The symmetric first order condition then is

\[
rf_e(y_c, e_c) = \frac{K'(e_c) - S_1(e_c, e_c) - S_2(e_c, e_c)}{V'(w_c)}.
\] (2.33)

Denote the associated behavioral function as \( e_c = \varphi^C_2(y_c) \). An immediate comparison of equations (2.32) and (2.33) reveals that the decentralized choice fails to take into account the externality/spillover of effort \( S_2(e_c, e_c) \). This may result in under- or over-input of effort for any given \( y_c \), depending on the sign of \( S_2(e_c, e_c) \). However, it would be too soon to draw any conclusions on equilibrium outcomes before considering the response of the parent. The parental investment \( y_c \) in any child is an implicit function of \( e_c \), \( y_c = \varphi_1(e_c) \), defined by

\[
rf_y(y_c, e_c) = \frac{U'(w_p - ny_c)}{aV'(w_c)},
\] (2.34)

in both cooperative and noncooperative cases. Notice that the symmetry between siblings has already been integrated into these \( \varphi \) functions. Thus, they do not characterize the full dynamics of strategic behaviors within the family as the \( \phi \)
functions do.

**Proposition 2.3.** In the symmetric model with sibling effects, assume positive spillovers of effort among siblings, \( S_2(e_c, e_c) > 0 \). Then, (i) for any given level of parental investment \( y_c \in [0, w_p] \), the cooperative effort input \( \phi_C^2(y_c) \) is always higher than the noncooperative input \( \phi_N^2(y_c) \), \( \phi_C^2(y_c) > \phi_N^2(y_c) \); and (ii) when \( \phi'_{1}(e_c)\phi_C'(y_c) < 1 \) for any interior \((e_c, y_c)\), cooperation among children leads to both higher child effort and parental investment in the equilibrium than the noncooperative equilibrium, \((y_C^c, e_C^c) > (y_N^c, e_N^c)\), under strong technological complementarity (2.10), while still higher effort, \( e_C^c > e_N^c \), but lower investment, \( y_C^c < y_N^c \), leaving the comparison of equilibrium outputs \( H_c \) undecided, under condition (2.15).

Part (i) of Proposition 2.3 is intuitive: potential coordination among children can internalize the externality of effort. Since \( S_2(e_c, e_c) > 0 \), condition (2.33) requires effort input until a point of lower return than (2.32).\(^{20}\) Part (ii) concerns equilibrium human capital production with cooperative and noncooperative choices among children. When the technological complementarity between investment and effort is strong, the tendency toward higher effort of the coordinated children also induces more parental investment (as shown in Figure 2.3). When the opposite holds, the parent invests less (as displayed in Figure 2.4), because consumption is more valuable. In the latter case, one might tend to conclude that the coordination should still give higher human capital output than the noncooperative choices even when

\(^{20}\)Rearranging (2.32) and (2.33) shows that the group marginal benefit of effort input \( rV'f_c + S_1 + S_2 \) is always higher than the private marginal benefit \( rV'f_c + S_1 \) for given \( y_c \). Further, the concavity of \( S(e_c, e_c) \) guarantees that the curve of group marginal benefit is downward sloping. It then only crosses the upward sloping marginal cost curve \( K'(e_c) \) once. The unique optimal effort choice also eliminates the difficulty of multiple solutions for coordination.
parental investment is lower. The argument would be that if it gave lower output, then the cooperative decision of higher effort would not be optimal. This reasoning is problematic because if the spillover of effort is so strong that the effect on group utility dominates, then a combination of higher effort and lower output under coordination can still be efficient for children. Put differently, when condition (2.15) holds, the implication of sibling interactions is rather uncertain for the equilibrium production of human capital.

On welfare, children must enjoy higher utility with coordination than in the noncooperative case; otherwise, they could simply use the noncooperative effort choices even during the cooperation. If the technological complementarity is not strong such that the parent invests less responding to higher effort, the parent is better off with cooperative children. This is because (i) the cooperation leads to higher child effort, which allows her to consume more, and (ii) the parent also enjoys the children’s higher utility due to altruism. If the complementarity is instead strong as in condition (2.10), then the parent will sacrifice more consumption for investment in children. Even with the higher utility of the children, the total effect of the cooperation on the parent’s welfare is still undecided in this later case.

Technically, Proposition 2.3 requires $\phi_1'(e_c)\phi_2'(y_c) < 1$ for part (ii). Under this condition, the children’s cooperative choice function $e_c = \phi_2(y_c)$ is steeper than the parent’s behavioral function $y_c = \phi_1(e_c)$ against $e_c$ when either (2.10) or (2.15) is in effect – both curves are upward sloping if the strong complementarity (2.10) is present while downward sloping if (2.15) holds; see Figures 2.3 and 2.4. Immediately, the two curves only intersect once, so the equilibrium with child coordination is
In a broad sense, Proposition 3 offers a direction for rethinking about social interactions. The implications of strategic interactions relative to coordination in games of strategic complementarities (Cooper and John, 1988) or, more generally, supermodular games (Milgrom and Roberts, 1990) have been extensively studied in the literature. Modeling family behaviors, this paper imposes another layer of interactions, parent-child interactions, upon the standard supermodular games among peers.
unique, as displayed in Figures 2.3 and 2.4. This condition might seem restrictive at first, but it is essentially the same as demanding that the aggregate utility of the parent and children should be concave, as in Proposition 2.2, which is a mild technical constraint.

In a broad sense, Proposition 2.3 offers an opportunity for rethinking social interactions. The implications of strategic interactions relative to coordination in games of strategic complementarities (Cooper and John, 1988) or, more generally, supermodular games (Milgrom and Roberts, 1990) have been extensively studied in the literature. Modeling family behaviors, this essay imposes another layer of interactions (i.e., parent-child interactions) upon the standard simultaneous game among peers (i.e., the inter-sibling game), and allows for the interplay between these forces. Proposition 2.3 simply characterizes the consequences of peer interactions in the full symmetric equilibrium. In many realistic situations, there are indeed endogenous feedback effects from background factors of the group to the behavior of group members in the game of social interactions. Existing models that treat contextual effects as exogenous seem incomplete for the purpose of understanding these more complicated group behaviors. The author believes that the current line of analysis is important for studying rich interactions within families in the process of human capital formation, and further, that it would be of interest in the theory to investigate complementarity games with principal players (as the parent in the inter-sibling game), although outside the scope of this essay.

For decision making in group contexts, it has also been known at least since Cooper and John (1988) that strategic complementarities among agents are neces-
sary for multiple equilibria to emerge. The argument can be adapted to the present setting and notations. Recall that child $i$’s reaction function $e^i_c = \phi^i_2(y^i_c, \bar{e}^{-i}_c)$ is determined by equation (2.31). Suppose that parental investment in each child is constant at $y_c$. Then, symmetric equilibria of sibling interactions occur at fixed points $e_c = \phi_2(y_c, e_c)$. Intuitively, if $\frac{d\phi_2(y^i_c, \bar{e}^{-i}_c)}{d\bar{e}^{-i}_c} < 0$ at $y^i_c = y_c$, the reaction function can only intersect the 45 degree line once and hence there must be a unique fixed point. However, under the strategic complementarity assumption (2.30),

$$\left. \frac{d\phi_2(y^i_c, \bar{e}^{-i}_c)}{d\bar{e}^{-i}_c} \right|_{y^i_c = y_c} = \frac{S_{12}(e^i_c, \bar{e}^{-i}_c)}{K'' - S_{11} - \tau (V'_{ee} + \tau V''_{ec})} > 0.$$  

Therefore, it is possible for function $\phi_2$ to have multiple fixed points so that multiple equilibria emerge in the inter-sibling game with fixed parental investment.

In spite of the complementary effect of sibling interactions, the parent must also react to the behavior of children in a full equilibrium. The derivative of function $\phi_2$ with respect to $\bar{e}^{-i}_c$ takes the form of

$$\frac{d\phi_2(y^i_c, \bar{e}^{-i}_c)}{d\bar{e}^{-i}_c} = \frac{\partial \phi_2(y^i_c, \bar{e}^{-i}_c)}{\partial y^i_c} \frac{\partial y^i_c}{\partial \bar{e}^{-i}_c} + \frac{\partial \phi_2(y^i_c, \bar{e}^{-i}_c)}{\partial \bar{e}^{-i}_c}.$$  

The fact $\frac{\partial \phi_2}{\partial \bar{e}^{-i}_c} > 0$ is not sufficient for concluding that $\frac{d\phi_2}{d\bar{e}^{-i}_c}$ here is positive. The equilibrium of sibling interactions has to account for the interplay with the higher layer parent-child interactions. This effect may eliminate the source of potential multiple equilibria in the behavior of children.

**Proposition 2.4.** In the symmetric setting of sibling interactions, if the parental choice
function \( y_i^c = \phi_1(e^i_c, y_{-i}^c) \) satisfies

\[
r (V'f_{ey} + rV''f_{e}f_{y}) \frac{\partial y_i^c}{\partial \bar{e}_{-i}^c} < -S_{12},
\]

then \( \frac{d\phi_2(y_i^c, \bar{e}_{-i}^c)}{de_{-i}^c} < 0 \) and there is a unique symmetric equilibrium.

Proposition 2.4 suggests that endogenous parental influences may dramatically change the dynamics of children’s strategic behavior. This again reminds one of the structure of intrafamily interactions, which includes the higher layer parent-child interactions upon the standard complementarity game among siblings/peers. Notice that the partial derivative \( \frac{\partial y_i^c}{\partial \bar{e}_{-i}^c} \) in (2.35) reflects the parent’s preferences and the production technology of human capital. There are two cases for a child to respond with lower effort to larger input of his siblings, i.e., \( \frac{d\phi_2}{de_{-i}^c} < 0 \). First, the parent compensates children working less hard (there are substitutions across children), \( \frac{\partial y_i^c}{\partial \bar{e}_{-i}^c} > 0 \), and the child tends to reduce effort when he receives more from the parent (the income effect of parental investment dominates the substitution effect), \( r (V'f_{ey} + rV''f_{e}f_{y}) < 0 \), as implied by condition (2.15). If the reduction in the incentive of effort input exceeds the increase \( S_{12} \) due to the strategic complementarity, the child’s input will be decreasing in his siblings’ average effort. The other case is that the parent rewards higher effort and accordingly reduces investment in child \( i \), \( \frac{\partial y_i^c}{\partial \bar{e}_{-i}^c} < 0 \); and this reduction of investment discourages the child’s effort input (the effect of parental investment on the productivity of effort dominates), \( r (V'f_{ey} + rV''f_{e}f_{y}) > 0 \), as given by condition (2.10). If the effect of siblings’ higher effort through this channel dominates the impact of the strategic
complementarity, the child will also exert less effort. In either case, the child’s effort is a downward-sloping function of average sibling effort, eliminating the possibility of multiple equilibria. In short, the interplay between parent-child interactions and endogenous sibling effects is important for a full analysis of family linkages in the development of human capital.

The Linear Quadratic Specification Revisited

This section extends the linear quadratic specification (2.16)-(2.20) to the multi-child setting considered in Section 2.3, and reexamines the implications of intrafamily interactions.

The preferences and technology are still the same as given by (2.16)-(2.20) except that all child-specific variables such as $H_{ci}$, $A_{ci}$, $y_{ci}$, $e_{ci}$, $a_{ci}$, $x_h$ and $w_{ci}$ are indexed with superscript $i$ to capture individual characteristics. The extended model also needs to specify $S(e_i, \bar{e}_i)$. Two important alternative formulations of group utility are proposed: one captures a pure conformity effect of the type studied by Bernheim (1994),

$$S(e_i^i, \bar{e}_i^{i-1}) = -\frac{1}{2} (e_i^i - \bar{e}_i^{i-1})^2, \quad J > 0; \quad (2.36)$$

and the other embodies a multiplicative interaction between individual and average effort,

$$S(e_i^i, \bar{e}_i^{i-1}) = Je_i^i \bar{e}_i^{i-1}, \quad 0 < J < 1. \quad (2.37)$$

This first specification (2.36) punishes deviations from the average effort of siblings, while the second (2.37) stands for the case of proportional spillovers. They are
equivalent when choices are binary, as shown in Brock and Durlauf (2001b), but not in this essay with continuous effort.

Denote the averages of parental investment, investment productivity, and effort input across \( n \) children as 
\[
\bar{y}_c = \frac{1}{n} \sum_{j=1}^{n} y_{jc}^i, \quad \bar{A}_c = \frac{1}{n} \sum_{j=1}^{n} A_{jc}^i, \quad \text{and} \quad \bar{e}_c = \frac{1}{n} \sum_{j=1}^{n} e_{jc}^i,
\]
respectively. Let \( \sigma \equiv \frac{a_1 (\eta_1 r \kappa)^2}{n \theta_2 + a \eta_2^2} \). Then, Proposition 2.5 follows.

**Proposition 2.5.** Assume an interior solution \( (\bar{y}_c^N, \bar{e}_c^N) \in (0, w_p) \times (0, \hat{e}_c) \) for the linear quadratic model specified by (2.16)-(2.20) and (2.36)-(2.37). Then, (i) the average parental investment \( \bar{y}_c \) is determined by
\[
\bar{y}_c = -\theta_1 + \theta_2 w_p + a \eta_1 r \bar{A}_c - \frac{n \theta_2 + a \eta_2}{n \theta_2 + a \eta_2} \bar{y}_c,
\]
and the average effort \( \bar{e}_c \) made by children in response to parental investment is
\[
\bar{e}_c = -\alpha + \eta_1 \kappa \bar{y}_c,
\]
if the group utility \( S(e_c^i, \bar{e}_c^{-i}) \) takes the form of (2.36), and
\[
\bar{e}_c = -\alpha + \eta_1 \kappa \bar{y}_c \frac{1}{1-J},
\]
if it is specified as in (2.37); and (ii) equilibrium averages \( (\bar{y}_c^N, \bar{e}_c^N) \) are unique, and the average choices \( \bar{y}_c^N \) and \( \bar{e}_c^N \) in the equilibrium are both independent of the degree \( J \) of sibling influences under (2.36), while both are increasing in \( J \) when the spillover effect among children dominates the effect of parent-child interactions, \( 1-J > \sigma \), or decreasing in \( J \) when the opposite \( \sigma > 1-J \) holds under (2.37).
Part (i) of Proposition 2.5 concerns the average behavioral responses of the parent and children. According to equation (2.38), average parental investment is proportionally increasing in family income $w_p$ and average productivity $\bar{A}_c$, but decreasing in the marginal benefit of consumption and the number of children. This latter effect captures the quantity-quality trade-off that the parent faces. In particular, when $n = 1$, equation (2.38) reduces to (2.21) in the single-child model. As for the children’s behavior, with either conformity effects or proportional spillovers, average effort is increasing in the productivity of effort, which depends on parental investment and the value of human capital, and naturally decreasing in the marginal cost of effort. However, an interesting difference between the two cases of group utility appears here: under proportional spillovers (2.37), the children’s average reaction to outside conditions is leveraged up by the multiplier $\frac{1}{1-J}$ as in equation (2.40), but conformity effects (2.36) do not have an analogous leverage mechanism, as evident from (2.39). Relative to the single-child (2.22), the multiplier $\frac{1}{1-J}$ in equation (2.40) captures a direct effect of sibling interactions.

Part (ii) of Proposition 2.5 presents further results on the equilibrium consequences of sibling interactions. Driven by conformity, sibling effects play no role in the average equilibrium choices, although as discussed below, they do have impacts on the choice distribution across children. Under proportional spillovers (2.37), an increase in $J$ leads to a higher leverage effect in the average behavior among children for any given parental investments; see equation (2.40). Taking into account the responses of the parent, however, this does not necessarily imply higher effort or investment in the equilibrium. As displayed in Figure 2.5, condition $\sigma > 1 - J$
Figure 2.5: Equilibrium Shifts Down when Sibling Influence Increases

Figure 2.6: Equilibrium Shifts Up when Sibling Influence Increases
ensures that the linear function \( \bar{y}_c(\bar{e}_c) \) (2.38) is steeper than the function \( \bar{e}_c(\bar{y}_c) \) implied by (2.40). An increase in \( J \) makes the function \( \bar{e}_c(\bar{y}_c) \) less upward-sloping, as \( \bar{e}_c(\bar{y}_c)' \) in Figure 2.5, which also lies below the original \( \bar{e}_c(\bar{y}_c) \) line. The aggregate effect is then shifting the equilibrium toward lower parental investment and child effort, \( (\bar{y}_c^N, \bar{e}_c^N) \). Of course, when \( 1 - J > \sigma \), a higher \( J \) of sibling influences indeed leverages more investment and effort in the equilibrium; the shift of the equilibrium is shown in Figure 2.6.

The notion of *social multipliers* is useful to conceptualize the effects of social interactions on group behaviors. Formally, a social multiplier measures the ratio of the effect on the average action caused by a change in a parameter to the effect on the average action that would occur if individual agents ignored the change in actions of their peers.\(^{21}\) Therefore, social interactions are a potential source of social multiplier effects; see Glaeser *et al* (2003) for an empirical example. However, the linear model constructed here shows that in a micro-founded framework, it is critical to distinguish the driving forces of social interactions in the determination of the social multiplier. For example, conformity is irrelevant for the average behavior of siblings. Furthermore, in the context of this essay with two layers of interactions, the social multiplier effects on the behavior of children have to be evaluated against the responses from the player at the higher layer – their parent, if one is interested in the implications of peer influences in a full equilibrium. When \( \sigma > 1 - J \), strong sibling influences may even lead to low effort choices because the

\(^{21}\)The social multiplier also appears in the literature defined as the ratio of the per capita response of the peer group to a change in the parameter that affects the entire peer group to the average response of an individual action to the same exogenous parameter that only affects that person; see Glaeser *et al* (2003) and Blume *et al* (2011a) for more discussion. This essay does not take this view.
parent’s behavior counteracts forcefully. When \( 1 - J > \sigma \), they do generate positive effects of sibling effort, but the social multiplier will include an adjustment factor \( \sigma \) to reflect parent-child interactions, given as \( \frac{1}{1-J-\sigma} \) in the full equilibrium.\(^{22}\) This multiplier is different from \( \frac{1}{1-J} \) in the behavioral function (2.40) of children when they treat parental investments as exogenous. Therefore, the interplay between parental influences and sibling effects may further amplify the impacts of peer behaviors, at least in the interesting case of \( 1 - J > \sigma \). In other words, the leverage effect of parent-child interactions found in the single-child model of Section 2.2 is again involved in the current setting of sibling interactions.

Besides the average behaviors examined in Proposition 2.5, individual decision functions are also informative. Substituting (2.38) into the parent’s first order conditions,

\[
y^i_c = -\theta_1 + \theta_2 w_p + a\eta_1 rA^i_c + \frac{\theta_2 \eta_1 r(A^i_c - \bar{A}_c)}{n\theta_2 + a\eta_2}.
\]

Parental investment in child \( i \) then contains her response to child \( i \)’s idiosyncratic productivity (the first term), partially determined by his effort, as well as a bonus to his extra productivity/effort relative to his siblings (the second term). In this linear model, the strong technological complementary condition (2.10) naturally holds, and the parent rewards children’s effort with greater parental investment.

\(^{22}\)This claim can be verified by solving for \( \bar{e}_c \) from equations (2.38) and (2.40).
order condition for the effort decision is

\[ e_i^c = \frac{-\alpha}{1 + \gamma} + \frac{J}{1 + \gamma} \bar{\epsilon}_c \bar{\eta}_i + \frac{\eta_1 \kappa}{1 + \gamma} \bar{y}_c. \]

It is straightforward from this expression that as long as \( J > 0 \), the child wants to keep up with his siblings. As emphasized, this is an important channel for endogenous transmission of human capital across siblings. He also increases effort when the parent invests more in him, and the extent of the increase is determined by the marginal value of effort \( r \kappa \) as well as the marginal utility of income \( \eta_1 \). Solve for his effort choice \( e_i^c \):

\[ e_i^c = \frac{n - 1}{n - 1 + \gamma} \eta_1 \kappa (y_i^c - \bar{y}_c) + \eta_1 \kappa \bar{y}_c - \alpha. \]

Child \( i \) not only reacts to average parental investment, but also to the parent’s investment in himself relative to his siblings. Proposition 2.5 has shown that pure conformity does not affect the average behavior of children. However, interestingly, it has impacts on individual decisions. According to this choice function, when the degree \( J \) of group influences increases, the role of individual attention (the first term) becomes weaker, and effort is determined more by the forces (the other terms) that are common to all children. An immediate result is that stronger conformity effects lead to more homogeneous behavior among children. Instead, under the group utility (2.37), child \( i \)’s optimal condition is given by

\[ e_i^c = -\alpha + J \bar{\epsilon}_c - \bar{\eta}_i \kappa \bar{y}_c. \]
and therefore the individual choice of child $i$ can be solved as

$$e^t_i = \frac{n - 1}{n - 1 + nJ} \eta_1 \kappa \tau (y^i_c - \bar{y}_c) + \frac{\eta_1 \kappa \tau \bar{y}_c - \alpha}{1 - J},$$

which again consists of an idiosyncratic component and a fixed component that is common to all children. In this case, the child’s reaction to outside conditions, captured by the fixed input, is leveraged up by the multiplier $\frac{1}{1 - J}$.

To summarize, the transmission of human capital within families is non-linear even in this highly stylized linear model, as similarly demonstrated in (2.24). Decision rules over parental investment and child effort are indeed linear under the current specification, but the interplay between them renders the equilibrium production of human capital involving each factor in a much more complicated way. The analysis again conveys the necessity of careful modeling of family linkages for any serious empirical investigation of inter- and intra-generational transmissions of human capital.

### 2.4 Parental Investments and Neighborhood Interactions

The acquisition of human capital occurs not just within families, but also in schools, neighborhoods, and social groups. Outside of families, parents usually impose no direct influence over the behavior of peers with whom their children interact, unless they are allowed to directly select peers for their children, for example, by choosing
a neighborhood in which to live. In this aspect, social interactions are distinguished from sibling effects, and there is a need to closely examine their interplay with intergenerational family linkages. This section considers the neighborhood as a reference group, which can be conveniently defined by geographic proximity, but can also be broadly interpreted as localities in a "social space" with some notion of proximity versus distance that is given content in Akerlof (1997). The analyses with exogenous and endogenous formation of neighborhoods are developed in turn.

**Interplay between Family Influences and Neighborhood Effects**

Neighborhoods affect their members in various ways. An important example is the local financing of public education in the United States, but this section will focus on contemporaneous behavioral influences among children in the same neighborhood. Similar to sibling effects, such endogenous interactions are usually driven by sociological and/or psychological factors. For instance, there may be an intrinsic desire to behave like certain others and hence the influences may be reciprocal. In terms of evidence, as a prominent example, Fryer and Torelli (2010) empirically document the existence of "acting white," a genuine issue concerning children’s popularity as judged by their peers, and Durlauf (2004) provides a critical overview on the empirical findings of endogenous neighborhood effects. To open the black box of neighborhood interactions, this section begins by considering the case in which the configuration of neighborhoods is exogenously given.

---

23For example, parents may prevent children from meeting with certain peers. This can be interpreted as choosing a reference group for their children in a social space.
There is a continuum of families $i$ with unit measure that reside in the same community $g$. Each family has one single parent and one child. All parents have the same preference,

$$U(e_p^i) + a \left[ V(w_c^i) + S(e_c^i, \bar{e}_c^g) - K(e_c^i) \right], \quad (2.41)$$

where $\bar{e}_c^g = \int e_c^i \, di$ is the neighborhood-wide average of child effort. Parents are endowed with different levels of wealth $w_p^i, \, i \in g$. Children are heterogeneous in ability and other characteristics, which are captured by their idiosyncratic production functions of human capital $f^i(y_c^i, e_c^i)$. Human capital still earns a constant market return $r$. And, children also have homogeneous preferences, implicitly specified in (2.41) as

$$V(w_c^i) + S(e_c^i, \bar{e}_c^g) - K(e_c^i), \quad (2.42)$$

where the social utility component $S(e_c^i, \bar{e}_c^g)$ depends on the average effort $\bar{e}_c^g$ in the community.\footnote{With a continuous set of family indices, there is no distinction between inclusive and exclusive averages.}

To derive conditions for the optimization problems of the parents and children, notice that they take neighborhood-wide average behaviors as given. It then immediately follows that the decision rules of the parents, $y_c^i = \phi_1^i(e_c^i)$ for all $i \in g$, are defined by

$$rf_{y}^i(y_c^i, e_c^i) = \frac{U'(w_p^i - y_c^i)}{aV'(w_c^i)}, \quad \text{for any } e_c^i \in [0, \hat{e}_c]. \quad (2.43)$$

Different from the case of sibling interactions, the function $y_c^i = \phi_1^i(e_c^i)$ here does
not depend on \( y^j_c \) if \( j \neq i \). Parental investments are not directly related between families in the community, and parents have no direct influence over the behavior of neighbor children. On the other hand, the reaction functions of the children, \( e^i_c = \phi^i_2(y^i_c, \bar{e}^g_c) \) for all \( i \in g \), are obtained from

\[
rf^i_c(y^i_c, e^i_c) = \frac{K'(e^i_c) - S_1(e^i_c, \bar{e}^g_c)}{V'(w^i_c)}, \quad \text{for any } y^i_c \in [0, w^i_p]. \tag{2.44}
\]

Interdependence among children’s effort choices builds upon their connections to the group average behavior. This relation becomes a bridge for the intragenerational transmission of human capital across families. Absent of endogenous social interactions, families would be self-contained and evolve in their own ways even living in the same community. In this sense, this essay provides a theory for neighborhood effects, and its unique feature is to allow these effects to also feed back upon intrafamily interactions.

A profile of investment and effort choices \((y^i_c, e^i_c)_{i \in g}\) is a Nash equilibrium of the social interactions model within a given neighborhood \( g \), if it satisfies both \( y^i_c = \phi^i_1(e^i_c) \) and \( e^i_c = \phi^i_2(y^i_c, \bar{e}^g_c) \), or equivalently (2.43) and (2.44), for all \( i \in g \). Appendix A.8 verifies that there exists a Nash equilibrium under the assumptions on preferences and technologies imposed by the present essay.

Furthermore, with identical families in the neighborhood, there is a unique symmetric equilibrium if \( \phi_{2y} \phi_{1e} > 1 \). In other words, in the symmetric environment, condition \( \phi_{2y} \phi_{1e} > 1 \) eliminates the possibility of multiple equilibria. To see the economic intuition, recall that strategic complementarities among peers are neces-
sary for multiple equilibria in coordination games. But the interplay with family linkages changes the incentive schemes of effort choices among children. When the average effort is lower in the neighborhood, a child’s direct response is less effort as well, because of the strategic complementary effect, $\phi_2 e > 0$. However, if his parent rewards effort heavily and parental investment increases productivity dramatically such that $\phi_2 y_1 e > 1$, this effect will offset the incentive to shirk. When this latter effect actually dominates, the child’s effort $e_i$ becomes a decreasing function of the neighborhood average $e^g$, which produces only one fixed point. This reasoning also explains the importance of intrafamily interactions for modeling neighborhood effects. It complements the existing literature on social interactions that usually concentrates on the inter-member relations only and isolates them from strategic connections between members and their background factors such as families.

To develop analytical results, the rest of the section deals with the linear specification as Section 2.3 that is adapted to the neighborhood interactions model. A straightforward calculation yields the decision rules of the parents,

$$y_i = -\theta_1 + \theta_2 w_p + a\eta_1 r A_i \theta_2 + a\eta_2,$$

(2.45)

for all $i \in g$, which is identical to the choice function (2.21) in the benchmark model. Parents tend to compensate their children’s effort or, more broadly, productivity, and wealthier parents proportionally invest more. The behavioral functions of the children are just the same as in the model of sibling interactions (see Section 2.3),
repeated here as
\[ e^i_c = \frac{-\alpha + \bar{J} + \bar{J} \tilde{e}^g_c + \eta_1 \bar{r} \bar{x}^i_c}{1 + \bar{J} - \sigma}, \quad (2.46) \]
\[ e^i_c = -\alpha + \bar{J} \tilde{e}^g_c + \eta_1 \bar{r} \bar{x}^i_c, \quad (2.47) \]

for all \( i \in g \), under the formulations of social utility (2.36) and (2.37), respectively. In this linear model of neighborhood interactions, the possible adjustment factor for the social multiplier is \( \sigma \equiv \frac{a |\eta_1 \bar{r} k|^2}{\bar{b}_2 + a \eta_2} \). To ease notation, let \( B^i_c = \frac{\eta_1 \bar{r} k}{\bar{b}_2 + a \eta_2} \left[ \left( \bar{\theta}_2 w^i_p - \bar{\theta}_1 \right) + a \eta_1 r \left( \rho a^i_c + \sigma H^i_p + \sum_h \beta_h \chi^i_h \right) \right] \). As in Section 2.2, \( B^i_c \) again can be interpreted as a comprehensive aggregation (adjusted to the parent’s behavior) of the background factors of child \( i \). This results in the following proposition.

**Proposition 2.6.** Given the reaction functions (2.45)-(2.47) in the linear model of neighborhood interactions, the effort choices of the children after taking into account the responses from their parents have the form of

\[ e^i_c = -\alpha + \bar{J} \tilde{e}^g_c + B^i_c, \quad (2.48) \]
\[ e^i_c = -\alpha + \bar{J} \tilde{e}^g_c + B^i_c - \bar{\sigma}, \quad (2.49) \]

among their peers in neighborhood \( g \), under social utility function (2.36) and (2.37), respectively. Therefore, the neighborhood-wide average of child effort is

\[ \bar{e}^g_c = \frac{-\alpha + \bar{B}}{1 - \bar{\sigma}}, \quad (2.50) \]
\[ \bar{e}^g_c = \frac{-\alpha + \bar{B}}{1 - \bar{J} - \bar{\sigma}}, \quad (2.51) \]
in the two cases. Then, the average effort \( \bar{e}_c^g \) is independent of \( J \) in (2.50) under social utility (2.36), and increasing in \( J \) if \( 1 - J > \sigma \) in equation (2.51) under social utility (2.37).

Proposition 2.6 characterizes the consequences of the interplay between family influences and social effects on the behavior of children. It is evident from expressions (2.48) and (2.49) that parental investments not only directly set the background incentives \( B_i, i \in g \), for their effort choices, but also scale up the dependence of individual decisions on the neighborhood average choice by the adjustment factor \( \sigma \) in the coefficients on \( \bar{e}_c^g \). Furthermore, equation (2.50) shows that the pure conformity effects among children play no role in average community behavior, although it affects the distribution of choices, as one can see from equation (2.46). In equation (2.51), with \( 0 < J < 1 \), the strategic complementarity amplifies the effect of any change in outside conditions, which leads to a social multiplier greater than one if \( 1 - J > \sigma \). In both cases, parental investments do not directly interact with each other, but are connected through the behavior of their children. The investments are proportional to child effort as given in (2.45) and hence exhibit the same patterns as effort choices.

As an analogy to the social multiplier, one may also define an intrafamily multiplier to summarize the effects of parent-child interactions on the behavior of children. Intuitively, it measures the ratio of the effect on the average action (e.g., effort choice) caused by a change in the outside conditions to the effect on the average action that would occur if individual children ignore the change in actions of their parents, i.e., when \( \sigma = 0 \). Then, the intrafamily multiplier is \( \frac{1}{1-\sigma} \) in expression (2.50) and \( \frac{1-J}{1-J-\sigma} \)
in (2.51), which are greater than one when $0 < \sigma < 1$ and $0 < \sigma < 1 - J$, respectively. The former multiplier is independent of the social parameter $J$. In the latter case, however, the intrafamily multiplier is affected by $J$. Therefore, under the social influence of proportional spillovers, the social multiplier and intrafamily multiplier work together to leverage the average effort and human capital formation in the community, as demonstrated in equation (2.51).

The analysis presented in this section describes how human capital is transmitted across generations in a nonlinear fashion, and how it spreads out in the neighborhood through endogenous social links. In this linear environment, intrafamily interactions potentially amplify the converging effect of social influences on the economic status of families in the same community, which is an important insight for understanding local economic and social dynamics.

Endogenous Neighborhood Formation

A complete investigation of neighborhood effects in human capital formation must account for how neighborhoods are formed in the presence of these effects. The rise in socioeconomic segregation observed in many Western countries indicates how community configurations evolve over time. With neighborhood choice available, parents respond to neighborhood effects through intrafamily interactions as well as direct selection of reference groups for their children. In this scenario, how does stratification emerge as an equilibrium outcome? How do intrafamily interactions interplay with neighborhood choices? This section tackles these issues by introducing endogenous family linkages and social interactions into the classic framework
of community formation developed by Benabou (1996).

Assume a continuum of families with unit measure, who may live in one of two communities, each of size 1/2 and denoted as A and B. Housing prices are \( \rho_g, g = A, B \), in the two communities. A family self-selects into a community by buying a house there from an absentee landlord. This is no mortgage market. These prices can be thought of as membership fees if non-residential reference groups are considered. There are two types of families, \( h \) and \( l \), which are defined by their wealth endowments \( w^h_p \) and \( w^l_p \), with \( w^h_p > w^l_p \). The proportions of the two types are \( m \) and \( 1 - m \). All parents are the same otherwise, and each family has only one child. Parents live two periods. In period 1, they choose communities to live in and pay housing prices. In period 2, they invest in their children, who also interact with peers in the same neighborhood; the setup of this second period is then just identical to what has been described in Section 2.4 with given community configurations. To focus on neighborhood decisions, assume that all children are ex ante identical in period 1.\(^{25}\)

**Proposition 2.7.** Housing prices are the same across neighborhoods, \( \rho_A = \rho_B \), and any neighborhood configuration is compatible with equilibrium, if (i) the social utility component \( S(e^c_i, \bar{e}^g) \) is absent from the utility function of children, or (ii) child effort is the only factor of human capital production, \( H^i_c = f^i_c(e^i_c) \).

\(^{25}\)To assume that wealth is the only source of heterogeneity among parents and that children are otherwise ex ante identical is largely for technical convenience. Wealth is certainly the most important indicator of economic conditions that matter for decisions concerning consumption, investment, and neighborhood. However, other attributes such as race, education, cultural background, and religion may also be relevant for neighborhood choice or even contribute to residential segregation in equilibrium. This chapter abstracts from these factors to concentrate on the roles of family linkages and neighborhood effects.
Proposition 2.7 basically asserts that family linkages and social interactions are both necessary for neighborhood choices to be relevant in the current framework. Absent of peer effects in their utility, children behave independently. Since parents do not derive utility from residential location, they only care about housing prices no matter whether they are rich or poor. This demand force equalizes housing prices across the communities to clear the market. When housing prices are equal, all parents are indifferent between neighborhoods. Any configuration then can emerge in an equilibrium. Similarly, when parent-child interactions are shut down, the parents have no incentive to choose neighbors either, simply because all children are ex ante identical. To choose peers for children is to choose their family backgrounds, but this matters only when the behavior of children is, to some degree, tied to their family status.

Back to the more realistic setting that this section starts with, equilibrium community formation is the outcome of the interplay between family influences and social effects on the behavior of children when they are both present. From the viewpoint of period 1 and with ex ante identical children, the optimized utility of a parent in period 2 is determined by her net wealth, which is the total endowment, predetermined by her type $t$, subtracting the housing price $\rho$ of the neighborhood in which she has chosen to live, as well as the configuration of the neighborhood, measured by its percentage $x$ of rich households. Denote this expected utility as $W(t, \rho, x)$. Then the equilibrium allocation of families across communities depends on how parents are willing to trade off the housing price $\rho$ against the neighborhood
quality \( x \). Specifically, let

\[
R(t, \rho, x) = \frac{d\rho'}{dx'} \bigg|_{W(t, \rho', x') = W(t, \rho, x)} = \frac{W_x(t, \rho, x)}{W_{\rho}(t, \rho, x)}
\]  

(2.52)

which is the slope of the family’s iso-utility curve. Benabou (1996) exploits this notion and proves the following important result.\(^{26}\)

**Lemma 2.1 (Benabou, 1996).** (i) If \( R(t, \rho, x) \) is increasing in \( t \) for all \( \rho \) and \( x \), the unique stable equilibrium is stratified. If \( m \leq 1/2 \), the rich type \( h \) families all live in one community, while if \( m \geq 1/2 \), the poor all live in a community. The symmetric equilibrium is unstable. (ii) If \( R(t, \rho, x) \) is decreasing in \( t \) for all \( \rho \) and \( x \), the unique equilibrium is completely integrated (i.e., symmetric), and it is stable.

The main intuition of Lemma 2.1 is that for stratification to be a stable equilibrium, the willingness to pay for a better neighborhood must be increasing in attribute \( t \) of parents. When this condition holds, any deviation of neighborhood composition away from symmetry will induce rich families to move to the wealthier community. The characterization of residential equilibrium to be presented in the rest of this section employs this tradeoff function \( R(t, \rho, x) \) and is built upon the result of Lemma 2.1.

Recall the results of Section 2.4 that in any given neighborhood \( g \), parental

\(^{26}\text{Becker and Murphy (2000) is another excellent exposition of the same idea but in a different format.}\)
investment in period 2 follows decision rule (2.43), updated as

\[ \text{rf}_y(y^i_c, e^i_c) = \frac{U'((\bar{w}^i_p - y^i_c))}{aV'((w^i_c))}, \tag{2.53} \]

and child effort is given exactly as (2.44), for all \( i \in g \). Here \( \bar{w}^i_p \equiv w^i_p - \rho_g \) is net wealth in period 2. Therefore, the endogenous choices of \( y^i_c \) and \( e^i_c \) are functions of state variables \( \bar{w}^i_p, g \) and \( x_g \) in neighborhood \( g \), written as \( y^i_c(g(\bar{w}^i_p, x_g)) \) and \( e^i_c(g(\bar{w}^i_p, x_g)) \). The neighborhood average effort \( \bar{e}^g \) depends on average net wealth \( \bar{w}^i_p \) and neighborhood composition \( x_g \) as \( \bar{e}^g(\bar{w}^i_p, x_g) \).

**Proposition 2.8.** In the present setting of endogenous neighborhood formation, for parents of type \( t \) living in neighborhood \( g \) with housing price \( \rho \) and composition measurement \( x \),

\[ R(t, \rho, x) = \frac{aS_2 \partial \bar{e}^g/\partial x}{U' - aS_2 \partial \bar{e}^g/\partial \rho}. \tag{2.54} \]

Suppose that \( \frac{\partial e^{i,g}_c}{\partial t} \) has the same sign, either \( \frac{\partial e^{i,g}_c}{\partial t} > 0 \) or \( < 0 \), for all \( i \in g \). Then, if

\[ S_{12} > \frac{S_2}{U' \frac{\partial \bar{e}^g/\partial t}{\partial t}}, \tag{2.55} \]

the only stable equilibrium configuration of families is stratified.

Proposition 2.8 gives a general characterization of the equilibrium of the model with endogenous neighborhood choice. Equation (2.54) is the marginal payment that parent \( t \) is willing to make for a marginal improvement in neighborhood affluence \( x \). Intuitively, the tradeoff is between the neighborhood effect and price
effect on average child effort transmitted through social and family linkages to the parent, \( aS_2 \partial \bar{e}_g / \partial x \) and \( aS_2 \partial \bar{e}_g / \partial \rho \), and the marginal value of consumption \( U' \) for the parent. Therefore, it essentially states that neighborhood decisions must jointly consider social effects and intrafamily interactions.

Proposition 2.8 further establishes necessary conditions for stratification to emerge as the only stable equilibrium. Take the case of \( \partial e_i, g / \partial t > 0 \) for all \( i \in g \) as an example. Notice that it is quite natural to have \( \partial U' / \partial t < 0 \) because wealthier parents usually tend to consume more. If there are spillovers of child effort, \( S_2 > 0 \), then \( \frac{s_2}{U} \frac{\partial U' / \partial t}{\partial e_i, g / \partial t} < 0 \) in (2.55). It is a standard argument that some degree of strategic complementarity, \( S_{12} > 0 \), is necessary for equilibrium stratification in classical models of community formation, such as Benabou (1996), in which vertical parent-child interactions and horizontal social interactions of children collapse to simple complementary effects among neighbors. The condition presented here is weaker than \( S_{12} > 0 \); in this example, \( S_{12} < 0 \) is compatible with equilibrium stratification as long as inequality (2.55) holds. The central idea is that in the current structure of two-layer interactions, the spillover effects \( S_2 \) have two channels to generate the dispersion in neighborhood evaluation between rich and poor families: the first is through the strategic complementarity in child behavior (children from distinct family backgrounds make different levels of effort and enjoy the spillover effects differently), while the second is to affect the parent’s utility directly (parents of distinct wealth levels evaluate the spillovers differently). For the second channel, observe \( \partial U' / \partial t \) in the right hand side of the inequality (2.55) and notice that it is absent from the standard models without parent-child interactions. The nature of the
strategic complementarity $S_{12} > 0$ determines that it is always a source of different values of the neighborhood to families. Without it, however, it is still possible that rich and poor families have different degrees of willingness to pay for a better neighborhood, as long as the effect of the second channel dominates. Proposition 2.8 essentially requires that rich families are willing to pay more as the aggregate result through the two channels. An analogous argument can be developed for the case of $\frac{\partial e_i^g}{\partial t} < 0$ for all $i \in g$. In short, adding family structure to a Benabou (1996) - style model so that children interact with their peers as well as their parents, Proposition 2.8 extends the understanding of socioeconomic stratification in a richer environment.

As a concrete example, consider the linear quadratic specification studied in Section 2.4 again, but allow for endogenous neighborhood choice. Under reasonable parameterizations that sustain an interior solution, equilibrium neighborhood composition turns out to be stratified in this simple environment. To facilitate the statement of the following proposition, denote $K \equiv \frac{\eta_1 r \kappa \theta_2 (w_p^h - w_p^l)}{\theta_2 + a \eta_2}$ and $H \equiv \frac{\eta_1 r \kappa \theta_2}{\theta_2 + a \eta_2}$.

Then, 

**Proposition 2.9.** In the linear specification of the endogenous neighborhood model, the

---

27 Condition $\frac{\partial e_i^g}{\partial t} > 0$ simply means that children from rich families exert more effort in their education. If parental investments are positively related to their wealth, this needs positive reactions of children to investment. This condition also implies that $\frac{\partial e_i^g}{\partial t} > 0$, because if rich children work harder, then average effort in a rich neighborhood must also be higher. The intuition for $\frac{\partial e_i^g}{\partial t} < 0$ is just the opposite, and Proposition 2.8 needs one of the two conditions to hold for any $i$ and $g$. 

tradeoff function (2.52) takes the form of

\[ R(t, \rho, x) = \frac{aS_2K}{(1 - \sigma)U' + aS_2H'} \]  

(2.56)

\[ R(t, \rho, x) = \frac{aS_2K}{(1 - \sigma)U' + aS_2H'} \]  

(2.57)

for the cases of social utility (2.36) and (2.37), respectively. The only stable equilibrium allocation of families across neighborhoods is stratified in both cases.

In this special linear case, rich parents are always willing to pay more for a better neighborhood than poor parents. Positive reactions of parents and children to each other reinforce neighborhood complementarities and result in stratification. Despite the same equilibrium allocation, Proposition 2.9 indicates that the payment willingness partly relies on the driving forces of social influences; conformity effects (2.36) and proportional spillovers (2.37) lead to different expressions of \( R(t, \rho, x) \). This may give different robustness properties of the equilibria in the two cases. Formulations (2.56) and (2.57) also explicitly show different leverage effects of the interplay between intrafamily interactions and social effects. In either case, notice that both \( S_2 \) and \( U' \) can be further specified in (2.56) and (2.57); details are found in Appendix A.12.

Intrafamily interactions are an important factor that shapes neighborhood effects. One has to consider how these forces interplay in the discussion of stratification or other community configurations as equilibrium self-selection. For example, this essay finds that a condition weaker than the strategic complementary is compatible
with equilibrium segregation in the present structure. Since family ties are largely ignored in Benabou (1996) and the related literature in urban and regional economics, this chapter complements the existing studies of neighborhood formation by modeling the relationship between family influence and social interactions.

2.5 Conclusion

The accumulation of human capital cannot occur without the cooperation of children. Their effort inputs are essential and determine the productivity of parental investment in their educational activities. This chapter develops a human capital theory which allows for strategic interactions between parents and children – investors and producers of human capital. Integrating competition and peer influence (e.g., conformity or spillover) among siblings into the analysis, the theory provides a characterization of family linkages along both intergenerational and intragenerational dimensions in the process of human capital formation. When children interact with their peers from other families, the interplay between family influence and social effects shapes the evolution of human capital and neighborhoods. Based on the linear quadratic specification of the model, additional implications are derived, some of which are conceptualized in the adjustment of social multipliers to intrafamily interactions. The unified framework established in this essay speaks to the important but distinct roles of family and society in the education of children. The themes involved are also likely to be relevant for a variety of urban and regional issues.
The theory can be generalized to incorporate many other important aspects of human capital. For example, it would be interesting to include heterogeneous human capital into the model. Becker (1993a) emphasizes the distinction between general human capital and specific knowledge, while Heckman et al (2006) argue the different effects of cognitive and noncognitive skills. Individuals have different initial attributes that determine comparative advantage in producing various types of human capital, which has implications for occupational choice. Parents respond to these heterogeneous incentive schemes, and this necessarily affects the interplay between family linkages and social interactions. Second, the current model does not have physical capital. Parents can transfer financial assets directly to their children in the form of bequests. When there are financial markets, the direct transfers also earn returns (e.g., interest) that are independent from the behavior of their children. Parents make additional tradeoffs between human capital and financial capital in response to the actions of their children. This complicates intrafamily interactions in an interesting way and potentially extends the theory in a fruitful direction. Third, with endogenous fertility decisions, how would social interactions change the quality-quantity tradeoff? In this situation, neighborhood effects not only interplay with intrafamily interactions but also family structures. Finally, human capital investments are subject to uncertainty, as is any investment in financial markets. There may be even an absence of bases for assigning probabilities over random events in the assessment of risk associated with investments. One then has to model how parents and children adjust to these uncertainties and/or ambiguities and examine the associated implications for social interactions. Introducing robustness
analysis into the present framework can potentially yield new insights on human capital dynamics and is of particular interest to the author. Open questions such as these have yet to be explored for a more complete understanding of human capital formation in theory.

As Becker (1993b) has convincingly argued, "While the economic approach to behavior builds on a theory of individual choice, it is not mainly concerned with individuals. It uses theory at the micro level as a powerful tool to derive implications at the group or macro level." This essay is written with a strong intention to provide a theoretical framework for further empirical study. The author does not pretend that the model developed so far can be readily taken to the data, although the linear quadratic specification is promising for empirical purposes. To construct an empirical counterpart of the analysis is indeed the first priority in the author's research agenda. To that end, identification has to be achieved before any concrete implications about behavior can be tested with survey and other data. Even in the simplest linear model, the identification of peer effects can fail, as first pointed out by Manski (1993) (see Blume et al (2011a) for a recent survey), let alone family influences and the interplay between these forces. This chapter closes by simply stating the author’s concern that careful empirical implementation is crucial for establishing evidence that family linkages and social interactions interplay in the formation of human capital.
3 MEASUREMENT ERROR AND POLICY EVALUATION IN THE FREQUENCY DOMAIN
3.1 Introduction

Measurement error is well understood to exist in most macroeconomic data. The fact that data are *ex post* revised from time to time indicates how common measurement error can be. For example, the U.S. Bureau of Economic Analysis monthly releases its updated measure of GDP and price indices of recent quarters. In 1983-2009, the average revision without regard to sign is about 1.1% for current-dollar quarterly GDP and 1.3% for real quarterly GDP [Fixler *et al* (2011, Page 12)]. Historical data are also *ex post* revised based on more complete information, as well as changes to methodology intended to more accurately reflect economic activities. About every five years, the U.S. government issues comprehensive revisions to past estimates of GDP. The latest July 2009 revisions reach back to 1929.\(^1\) This is of course hardly unique to the United States. One striking example is China’s 2005 GDP revision. In light of the country’s first nationwide economic census, China’s statistics bureau revised its measure of 2004 national GDP upward by 16.8%. This substantial revision moved China above Italy as the sixth-largest economy in the world in 2004. For variables that are defined as the differences between actual and baseline values, the measurement problems become even acuter under structural change when baseline values may vary unpredictably. As an example, Orphanides *et al* (2000) and Orphanides and van Norden (2002) empirically documented errors in the measurement of the output gap for the U.S. economy, a part of which arise from the unobservable baseline of potential GDP. This type of data noise is serious when

\(^1\)The Federal Reserve Bank of Philadelphia maintains a real-time dataset of the U.S. economy which consists of 23 quarterly macroeconomic variables from 1965 to the present and includes historical revisions to these variables in great details [Croushore and Stark (2001)].
it is difficult to distinguish temporary shocks from permanent changes.

Monetary policy must be made in real time and so necessarily uses noisy data. Standard policy rules represent mappings from current and past economic conditions to monetary policy instruments such as the money supply or interest rates. For example, the famous Taylor (1993) rule is a linear mapping of observations of inflation and the output gap to the federal funds rate. At the time of the interest rate choice, the rule data available are therefore preliminary and with considerable measurement noise. In short, policymakers must live with and account for measurement error. But how should measurement error affect policy choices? How does measurement error affect the robustness properties of policy rules when the knowledge about fundamental economic structures is imperfect? Does this information constraint justify policy cautiousness, and if it does, how? Considering the possibly large welfare costs and long-lasting economic consequences associated with inflation and economic fluctuations, these questions are important in assessing alternative monetary policies.

To address these issues, this chapter contributes to the policy evaluation literature by investigating the implications of measurement error for the design of stabilization policy rules in the frequency domain. As pointed out by Orphanides (2003), this informational limitation on the true macroeconomic variables facing policymakers has been noticed in policy analysis since at least Friedman (1947). Orphanides’ (2001, 2003) own work largely reignites research interest in monetary policy evaluation with noisy information; recent contributions include Aoki (2003), Coenen et al (2005), Croushore and Evans (2006), Molodtsova et al (2008), and Or-
phanides and Williams (2002) among others. Many of these studies focus on the use of real-time data and evaluate the performance of real-time policies against ex post revised data.\(^2\) However, there has yet to be any systematic examination of the role of measurement error on policy choice in the frequency domain. Frequency domain approaches have been part of macroeconomic analysis for several decades – Hansen and Sargent (1980), Whiteman (1985, 1986), and Sargent (1987) are standard examples – and have recently experienced a resurgence in the context of policy evaluation [e.g., Brock and Durlauf (2005), Brock et al (2008a) and Hansen and Sargent (2008, Chapter 8)]. This chapter develops strategies to characterize frequency-specific performance of alternative policy rules in exposure to data noise.

There are significant reasons why frequency-specific analysis is important for policy evaluation. First, a full characterization of policy effects frequency by frequency is informative to policymakers. In the frequency domain, stabilization policy may be understood as determining the spectral density matrix of the state variables concerned. A full understanding of the effects of a policy rule requires evaluating how cycles at all frequencies are reshaped by the policy. When variances at some frequencies have greater social welfare costs than variances at other frequencies, it is necessary to know frequency-specific performance of alternative policies in order to make sound policy recommendations. This differential weighting of variance by frequency will occur, for example, when the social loss function involves non-time-separable preferences [Otrok (2001)]. In the case of

\(^2\)For example, Orphanides et al (2000) and Orphanides and van Norden (2002) showed that measurement errors are significantly large in real-time estimates of the output gap so as to render the estimates highly unreliable as guides to policymaking if data noise is not appropriately accounted for.
committee policymaking, it is possible that some committee members care more about performance at low frequencies while others care more about performance at high frequencies, hence this information is needed to allow for successful group decisionmaking.

Second, a number of properties of stabilization policies can really only be understood in the frequency domain. Policies that perform well at all frequencies are naturally appealing to policymakers regardless of their preferences. However, it turns out that such policies do not exist. Even if a policy reduces aggregate variance relative to some baseline, for the framework I study this will necessitate increasing variance at some frequencies in exchange for reducing variance at others. These tradeoffs are known as design limits. They were first identified by Bode (1945) in the engineering literature of linear system control and were introduced into the study of feedback policy rules in macroeconomics by Brock and Durlauf (2004, 2005) with extension to the vector case with forward-looking elements developed by Brock et al (2008b). These design limits are sufficiently complicated in the time domain as to render use impractical outside of the frequency domain.

In this direction, this essay contributes to the existing literature by studying design limits in the presence of measurement error. As such, the essay extends the study of design limits to the empirically salient case in which a policymaker is ignorant the true state of the economy due to measurement imperfections. In the linear feedback control system, the presence of measurement error creates new design limits than those that have been identified. Intuitively, a feedback policy rule introduces undesirable side noise into the system responding to noisy data,
when it exerts influences on the state variable to stabilize the economy. And when
the responses are aggressive, the side noise effects are also strong. Therefore, good
variance-reducing control has to be traded off against suppression of side noise.
Put differently, facing noisy data the policymaker has to make tradeoffs across
frequencies as well as between the channels of stabilizing control effects and side
noise effects. These constraints are summarized by two concepts – Bode’s (1945)
integral formula and the complementary principle [Skogestad and Postlethwaite
(1996, Chapter 5)], and are thus amenable to analytical treatment. This essay is the
first to put them to work in the practice of policy evaluation.

I further examine the effects of measurement error on policy design in the
presence of model uncertainty. When policymakers are uncertain about the true
model of the economy, the conventional wisdom is that policy reactions to the
observed state variables should be less aggressive; see Brainard (1967) for a clas-
sic example with parameter uncertainty and a known probability density on the
parameters, and Giannoni (2002) for a recent case of parameter uncertainty but
unknown distribution. However, little has been known about policy behavior when
the observations are also noisy. Allowing for both sources of uncertainty – mea-
surement error and model uncertainty, the analysis presented in this chapter sheds
light on robustness of efficient policy rules that recognize the presence of data noise.
Following the literature on robustness pioneered by Hansen and Sargent (2001,
2003, 2008), I assume that the true model is local to a baseline model and employ the
minimax decision criterion to evaluate policies against potential deviations from
the baseline. The minimax criterion is appealing in this context because deviations
from the baseline are, given the assumption of local model uncertainty, empirically indistinguishable so that, unlike in Brainard (1967), one has no basis outside of prior beliefs for assigning probabilities to alternative models. This essay models potential model misspecifications as variations of the spectral density function of the state variable, which conceptually distinguishes from Giannoni’s (2002) parameter uncertainty. This nonparametric approach to model uncertainty allows one to apply design limit results in a straightforward fashion to construct robust feedback policies, *i.e.* policies that work well regardless of which element of the model space is the true model.

Concretely, this essay focuses on linear feedback control rules in a one-equation backwards-looking model with single control input.³ The control is chosen by a policymaker to stabilize the economy in the sense of minimizing the variance of the state variable of interest. I show how measurement error, which produces a stochastically perturbed optimization problem with noisy control in the time domain, can be represented as a deterministic perturbation in the frequency domain. Further, each feedback rule can be associated with a sensitivity function in the frequency domain which describes how the policy shapes the spectral density of the state variable of interest. I characterize how the sensitivity function behaves differently in the presence and absence of data noise to show the frequency-specific effects of measurement error on optimal and robust policies.

³Central banks may also need to account for the effects of policies on expectations. Thus, it is important to address measurement error issues in a forward-looking framework as well. I leave this problem for future research. However, as Fuhrer (1997) has shown, expectations of future prices are empirically unimportant in explaining price and inflation behavior. In light of such evidence, the attention of this essay to backward looking models is natural and also relevant.
Different policy scenarios are considered and compared. First, the policymaker simply ignores the measurement issue and naively adopts the optimal policy for the standard noise-free problem. I show that this leads to a policy rule that is excessively aggressive. Failing to acknowledge data noise causes undesirable side effects with an activist policy. Second, I consider the case in which the policymaker is aware of measurement error and uses this information to adjust to an efficient policy rule design. The policymaker is more cautious in this case, but the adjustments in optimal policy relative to the noise-free case differ across frequencies. Third, I examine the situation in which the policymaker filters the noisy data to reduce measurement inaccuracy and applies the noise-free optimal rule to the filtered data. The Wiener filter used in this scenario effectively accounts for measurement error, as shown in the numerical exercises. Fourth, following Brock and Durlauf (2005), I perform local robustness analysis by approximating the robust policy solution with a small level of model uncertainty around the equilibrium solution to the baseline model using the minimax criterion. This worst-case study is modeled as a zero-sum game by introducing an adversarial agent who selects a model from a small neighborhood of the baseline model to maximize the loss function against the policymaker. I show the differences between the robust and standard solutions to illustrate the interaction between concerns over measurement noise and model uncertainty in policymaking.

Finally, for numerical implementation and more specific conclusions on monetary policy, I apply the theoretical analysis to an AR(1) monetary model, which is a variant of the two-equation Keynesian model. I start with a simple rule in
which the control only depends on the current state; lagged terms are excluded. This is interesting because the effects of measurement error can be characterized by one single coefficient parameter in this simple case. Also, simple instrument rules have received much attention in related literature. The optimal policy adjusted for the measurement turns out to be less responsive than a naive policy that assumes measurement is exact. In frequency domain, adjusting the policy design or filtering the data to account for measurement error, which is assumed to be a white noise process, the policymaker generally becomes less responsive to raw data observations; yet, the data filtering method outperforms the policy adjusting approach. Measurement error has little impact at low frequencies but results in more cautious policy reaction at high frequencies, and even may lead to more active control at medium frequencies. Without measurement error, model uncertainty has similar policy implications; it has little effect at low frequencies, reduces the strength of control at high frequencies, and increases control at medium frequencies. Therefore, Brainard’s (1967) intuition that model uncertainty leads to less effective policies follows in the sense of the total effect over the whole frequency domain, but fails at medium frequencies where the robust control is actually becoming more active. Introducing measurement error, however, the policymaker will reduce his reaction to model uncertainty, especially at medium and high frequencies. In other words, facing various types of uncertainty the policymaker’s reaction to one type of uncertainty is weakened by his attention to another type, although such effects differ across frequencies.
3.2 General Framework

To develop a general framework for policy analysis in the presence of measurement error, I start by formulating the policymaker’s problem in the frequency domain. Then, I turn to consider scenarios in which the policymaker reacts to mismeasurements differently, and evaluate policy performance in each case.

Policy Evaluation in the Frequency Domain

Suppose that the economy is governed by a simple scalar version of a backwards-looking dynamic system and the policymaker wishes to stabilize the economy in the sense of minimizing the variance of the economic variable of concern. The system is described as

\[ x_t = A(L)x_{t-1} + B(L)u_{t-1} + W(L)\varepsilon_t, \]  

(3.1)

where \( x_t \) is the unobserved random variable of interest with measure \( x_t^* \), both of which have zero means, and \( \{\varepsilon_t\} \) is a process of fundamental innovations with variance \( \sigma_\varepsilon^2 \). The control \( u_{t-1} \) is restricted to be a linear feedback rule and only feeds back to current and past observations, rather than underlying true realizations, of variable \( x_t \):

\[ u_{t-1} = -F(L)x_{t-1}^*. \]  

(3.2)
When there is no measurement error in the data, \( x^*_t = x_t \) for any \( t \). Otherwise, I model measurement error as an additive process \( \{n_t\} \) to \( \{x_t\} \), that is,

\[
x^*_t = x_t + n_t, \text{ with } n_t = D(L) \eta_t.
\] (3.3)

Here, \( \{\eta_t\} \) is another process of fundamental innovations orthogonal to \( \{\varepsilon_t\} \), \( \mathbb{E}[\eta_t \varepsilon_{t-k}] = 0, \forall k \), and hence uncorrelated with \( \{x_t\} \). The variance of \( \eta_t \) is \( \sigma^2_\eta \). The linear representation of \( n_t \) does not require that that measurement error is white noise, as is usually assumed without justification. Thus, \( \{n_t\} \) may be autocorrelated. The objective of stabilization policy design is to minimize the unconditional variance of \( x_t \),

\[
V(x) = \mathbb{E}[x_t^2],
\] (3.4)

by choosing an optimal rule \( u_{t-1} \). I consider this very simple loss function to focus on the study of measurement error effects, but the analysis can be extended to the case of more complicated preferences without difficulty, especially when I move to solve the problem in the frequency domain. In the presence of measurement error, the challenge for the policymaker is that he only observes \( x^*_t \) but needs to stabilize \( x_t \). If there is no such data limitation, he simply solves a standard feedback control problem with a single input and a single output.

The lag polynomials \( A(L), B(L), W(L), F(L), \) and \( D(L) \) are assumed to have only nonnegative-power terms. This one-sided specification rules out any forward-looking element in the model, and allows us to avoid complexities that arise from the formation of expectations. Some studies, for example, Fuhrer (1997), have
shown that expectations elements are not empirically important in explaining the
dynamics of inflation and output. Previous work on measurement error such as
Orphanides (2001, 2003) also uses backwards-looking models.

In the time domain, when the control is set to zero, \( F(L) = 0 \), the uncontrolled state variable \( x_{t}^{nc} \) has a moving-average representation as

\[
x_{t}^{nc} = (1 - A(L)L)^{-1} W(L) \varepsilon_{t}.
\]

(3.5)

I assume that the uncontrolled model (3.5) itself is stationary. This is without loss of generality because the unit root can be removed from a nonstationary model before we take it to consider the policy question. On the other hand, when a noisy feedback rule \( F(L) \) as specified by (3.2) is applied, the controlled state variable \( x_{t}^{c} \) takes the form of

\[
x_{t}^{c} = -T(L)n_{t} + S(L)x_{t}^{nc},
\]

(3.6)

where \( S(L) \equiv \frac{1}{1 + G(L)} \) and \( T(L) \equiv \frac{G(L)}{1 + G(L)} \), with \( G(L) \equiv [1 - A(L)L]^{-1}[B(L)F(L)L] \). Stabilization imposes the invertibility of \( G(L) \) and \( 1 + G(L) \). It is clear from equation (3.6) that there are two sources of volatility for the controlled state variable \( x_{t}^{c} \): original system and measurement error. Specifically, there are irreducible stochastic components in the state variable, and the feedback control causes undesirable side noise effects when responding to noisy data.

I use the following notations to facilitate work in the frequency domain. The Fourier transform of the coefficients of an arbitrary lag polynomial \( C(L) \),

\[
C(e^{-i\omega}) = \sum_{j=-\infty}^{\infty} C_{j} e^{-ij\omega},
\]

is denoted as \( C(\omega) \), \( \omega \in [-\pi, \pi] \). I define the sensitivity function
S(ω) and the complementary sensitivity function T(ω) in the frequency domain associated with a given feedback rule F(L) as the Fourier transforms of S(L) and T(L), respectively. Thus,

\[ S(\omega) = \frac{1}{1 + G(\omega)} \quad \text{and} \quad T(\omega) = \frac{G(\omega)}{1 + G(\omega)}, \]  

(3.7)

where \( G(\omega) = [1 - A(e^{-i\omega})e^{-i\omega}]^{-1} [B(e^{-i\omega})F(e^{-i\omega})e^{-i\omega}] \). In general, both S(ω) and T(ω) are complex functions. They sum up to one at each frequency in \([-\pi, \pi]\),

\[ T(\omega) + S(\omega) = 1. \]  

(3.8)

This is known as the complementary principle [Skogestad and Postlethwaite (1996, Chapter 5)].

To specify a feedback rule F(L) in the time domain is equivalent to characterizing the associated S(ω) and T(ω) in the frequency domain. Notice that all elements except F(L) in the expressions of S(L) and T(L) are exogenously determined. Thus, a chosen feedback rule F(L) will determine S(L) and T(L), and in turn determine their Fourier transforms S(ω) and T(ω). On the other hand, the coefficients of the lag polynomials S(L) and T(L), and therefore those of F(L), can be discovered from S(ω) and T(ω) by the Fourier recovery formula [Priestley (1981, Chapter 4)]. Therefore, one can always infer the feature of a linear feedback rule by reverse engineering, once clear about the behavior of its sensitivity function. For the rest of the chapter, instead of deriving feedback policy rules directly, I will focus on the characterization of their sensitivity functions to study the frequency-specific
implications of measurement error.

Every second-order stationary stochastic process admits a *spectral density function* that describes how the variance of a time series is distributed with frequency. Let \( f_{x^c}(\omega) \), \( f_{x^{nc}}(\omega) \), and \( f_n(\omega) \) be the spectral density functions of controlled state variable \( x^c_t \), uncontrolled state variable \( x^{nc}_t \), and measurement error \( n_t \). Given the expression (3.6) of \( x^c \), I can represent its spectral density function \( f_{x^c}(\omega) \) as

\[
f_{x^c}(\omega) = |T(\omega)|^2 f_n(\omega) + |S(\omega)|^2 f_{x^{nc}}(\omega),
\]

(3.9)

where \(|·|\) stands for complex modulus, \(|S(\omega)|^2 = S(\omega)S^*(\omega)\) and \(|T(\omega)|^2 = T(\omega)T^*(\omega)\), with superscript * referring to the complex conjugate. By applying the spectral representation theorem to (3.6), this result follows immediately from the fact that \( \eta_t \) and \( \varepsilon_t \) are independent from each other [Priestley (1981, Chapter 4)].

Equation (3.9) conveys the nature of the policy design problem with noisy data. The variance is just the integral of a variable’s spectral density function over \([-\pi, \pi]\). Functions \( f_{x^{nc}}(\omega) \) and \( f_n(\omega) \) represent the allocation of variance across frequencies for the uncontrolled state variable \( x^{nc}_t \) and measurement error \( n_t \), respectively. The sensitivity function \( S(\omega) \) specifies how the control redistributes the variance of

---

4 A quick derivation: Let \( dZ(\omega) \) be the increment in power, under a Fourier integral, over an infinitesimal interval \( d\omega \), then \( dZ_{x^c}(\omega) = S(\omega)dZ_{x^{nc}}(\omega) - T(\omega)dZ_n(\omega) \). Since the independence, \( E[dZ_{x^{nc}}(\omega)dZ_n(\omega)] = 0 \). It follows that

\[
f_{x^c}(\omega) = \frac{E[dZ_{x^c}(\omega)]^2}{d\omega} = \frac{E[S(\omega)dZ_{x^{nc}}(\omega) - T(\omega)dZ_n(\omega)]^2}{d\omega} = |T(\omega)|^2 f_n(\omega) + |S(\omega)|^2 f_{x^{nc}}(\omega).
\]

This expression also holds as a special case of the cross spectrum formula [see Sargent (1987, Page 248)].
the state variable frequency-by-frequency by reshaping $f_{xnc}(\omega)$. Since the control responds to noisy observations, it introduces additional noise into the system. The complimentary sensitivity function $T(\omega)$ captures the size of this side effect at each frequency, which acts on the variance distribution $f_n(\omega)$ of measurement error. At each frequency, the total effect from the two channels is then summed up as given in (3.9). The variance of the state variable under control is distributed with frequency according to $f_{xn}(\omega)$.

An ideal control rule would be one that lowers variance at every frequency to zero. However, limitation in control design make it impossible to reduce power uniformly over the whole frequency domain, even without data noise effects. This is known as Bode’s (1945) constraint; for a backwards-looking system, the sensitivity function associated with a feedback rule is subject to the integral restriction

$$\int_{-\pi}^{\pi} \ln(|S(\omega)|^2) \, d\omega = K_s, \text{ with } K_s \geq 0.$$ (3.10)

If the uncontrolled system is stationary as assumed, then $K_s = 0$; otherwise, $K_s > 0$. Notice that if $|S(\omega)| < 1$ for all $\omega \in [-\pi, \pi]$, then $K_s \geq 0$ cannot hold in equation (3.10). Therefore, reductions of variance at some frequencies for the state variable induce increases in its variance at other frequencies. Policy design has to make tradeoffs among the magnitudes of different frequency-specific variance contributions. A simple proof of this constraint (3.10) on sensitivity function can be found in Section B.1 of the Appendix.\(^5\)

\(^5\)Since $S(\omega)$ and $T(\omega)$ sums up to 1 at each frequency, there exists a similar Bode’s integral constraint and interpretation for $T(\omega)$. A recent version of proof for Bode’s integral constraint on
Now let me state the policy design problem in the frequency domain. Recall that \( \mathbb{E}[x_t^2] = \int_{-\pi}^{\pi} f_{x_c}(\omega) d\omega \) in the loss function (3.4), which can be conveniently generalized to more complicated – for example, non-time-separable – preferences using a general weighting function of variance by frequency. Notice that \( f_{x_c}(\omega) \) is given by (3.9). Then, I first consider the standard noise-free control problem as a benchmark. Since \( f_n(\omega) = 0 \) for all \( \omega \in [-\pi, \pi] \), \( T(\omega) \) becomes irrelevant in the objective function. The policymaker’s problem is just to choose a sensitivity function \( S(\omega) \), which is equivalent to his choice of a feedback rule \( F(L) \) in the time domain, to solve

\[
\min_{S(\omega)} \int_{-\pi}^{\pi} [S(\omega)]^2 f_{x_n}(\omega) \, d\omega,
\]

subject to Bode’s constraint (3.10) with \( K_s = 0 \). Following Brock and Durlauf’s (2005) argument, the optimal feedback rule is to reduce the system to a white noise process. In this benchmark case, the controlled state variable is

\[
x_t^B = [1 - (A(L) - B(L)F(L))L^{-1}W(L)\varepsilon_t].
\]

Then, the optimal feedback rule \( F^B(L) \) is

\[
F^B(L) = B(L)^{-1} \left( [W(L)L^{-1}]_+ + A(L) \right),
\]

where \( [\cdot]_+ \) is the annihilation operator. The sensitivity function \( S^B(\omega) \) associated...
with this benchmark control satisfies

\[ \left| S^B(\omega) \right|^2 = \frac{\sigma^2}{2\pi f_{xnc}(\omega)}. \]  

(3.14)

In the frequency domain, the benchmark control targets the constant spectral density function of fundamental innovations for the controlled state variable and achieves that by reallocation of variance across frequencies.

Back to the interesting case with noisy data, \( T(\omega) \) enters into equation (3.9) and hence the objective function (3.4). The policymaker chooses both \( S(\omega) \) and \( T(\omega) \) to minimize the loss function,

\[
\min_{S(\omega), T(\omega)} \int_{-\pi}^{\pi} \left[ |T(\omega)|^2 f_n(\omega) + |S(\omega)|^2 f_{xnc}(\omega) \right] d\omega,
\]

subject to the complementarity condition (3.8) and Bode’s constraint (3.10). As noted above, in the presence of measurement error, any control introduces side noise into the system, which is captured by the complementary sensitivity function. The policymaker then has to balance the conflicting effects between stabilizing control and side noise. Put differently, a compromise has to be made in control design when information is noisy; good control and disturbance rejection must be traded off against suppression of side noise process. This is why the complementary principle becomes important in this context. I will consider scenarios in which the policymaker deals with measurement error differently in the next subsections.

Let me close this subsection by a discussion on the notion of control aggressiveness.
It is natural to expect that measurement error may change the aggressiveness of the control used by the policymaker in response to current and past observations. In the time domain, the coefficients of the feedback rule $F(L)$ reflect responses to the data. Focusing on nonparametric sensitivity functions in the frequency domain, however, there is not such an obvious measure. I therefore propose the following notion of aggressiveness $AG$:

$$AG \equiv \left[ \int_{-\pi}^{\pi} (1 - |S(\omega)|)^2 d\omega \right]^{\frac{1}{2}} . \quad (3.16)$$

In $L^2$ norm, it measures how close the modulus function $|S(\omega)|$ to the constant function 1. Notice that $|S(\omega)| = 1$ means that the control is inactive at $\omega$, while the more $|S(\omega)|$ deviates from 1, the more powerful the control is at this frequency either shifting up or down the uncontrolled spectral density function. $AG$ measures control aggressiveness by the total deviation of $|S(\omega)|$ from 1 over the whole interval $[-\pi, \pi]$. Furthermore, similar to the benchmark case, I show that the policymaker’s behavior when he faces noisy data can still be interpreted as targeting some constant level $\lambda$ for the spectral density function of the state variable, although this may not be actually achieved due to side noise effects. In the same environment, the higher the level $\lambda$ the policymaker targets, the less aggressive the control would be.\(^6\) Hence, I consider the target level for the controlled spectral density function as another way for understanding control aggressiveness. In the numerical exercises, I will use both notions.

\(^6\)If the original systems are different, then the same target level $\lambda$ for the spectral density may reflect different degrees of control aggressiveness in the $AG$ measure.
**Naive Policy Rule**

A policymaker’s decision depends on his knowledge of measurement error. Let me first examine the case in which he is not aware of measurement issues and simply considers the observations as the true realizations of the state variable.

Since the policymaker knows the model, then he just naively adopts the optimal feedback rule for the noise-free system $F^B(L)$ from (3.13) even in the presence of measurement error. That is,

$$ u_{t-1}^N = -F^B(L)x_t^* . $$  \hspace{1cm} (3.17)

The associated sensitivity function $S^N(\omega)$ is still the same as (3.14),

$$ |S^N(\omega)|^2 = \frac{\sigma^2_\varepsilon}{2\pi f_{\lambda x^\epsilon}(\omega)}. $$  \hspace{1cm} (3.18)

I refer to $u_{t-1}^N$ or $S^N(\omega)$ as the *naive policy rule* in Orphanides(2003) terminology. By using this naive rule, the policymaker still targets the constant spectral density function of fundamental innovations for the state variable:

$$ \lambda^N = \frac{\sigma^2_\varepsilon}{2\pi}. $$  \hspace{1cm} (3.19)

However, the naive policy rule is inefficient with noisy data. The state variable under the control of the naive rule obeys

$$ x_t^N = \varepsilon_t + W(L)^{-1}[1 - W(L) - A(L)L]D(L)\eta_t, $$  \hspace{1cm} (3.20)
which admits spectral density function

\[ f_{xN}(\omega) = \frac{\sigma^2}{2\pi} + |D(\omega)|^2 |1 - S_N(\omega)|^2 \frac{\sigma^2}{2\pi}. \] (3.21)

Therefore, the naive policy rule is unable to achieve its target spectral density as in the noise-free benchmark; rather, it induces extra volatility at each frequency because of control noise effects.

The performance of the naive policy rule illustrates the consequences of ignoring potential measurement error in policymaking. The noise-free optimal policy rule will not be able to reduce the system to the white noise as it would when measurement error was absent. As a special case, even when the measurement error process is white noise, \( D(\omega) = 1, \forall \omega \in [-\pi, \pi] \), the system under the control of the naive policy rule is still not a white noise process; this can be verified by observing that \( |1 - S(\omega)| \) is not constant over frequencies in expression (3.21). The second term of (3.21) describes the frequency-specific effects of measurement error when the policymaker uses the naive policy rule. As shown in the numerical exercises in Section 3.4, either in terms of the AG measure or spectral target, the naive policy rule is too active relative to the optimal policy rule that solves (3.15).

**Optimal Policy Rule with Nonfiltered Data**

Suppose now that the policymaker recognizes the presence of measurement error as well as the error-generating process. He does not filter the data, but chooses an optimal control to account for measurement error and stabilize the economy. The
policymaker still restricts himself within the set of linear feedback rules as specified by (3.2). I will study the properties of this optimal policy rule with nonfiltered data by characterizing its sensitivity function $S^E(\omega)$.

The policymaker’s optimization problem is posed as (3.15) in the frequency domain. From (3.8),

$$|T(\omega)|^2 = 1 + |S(\omega)|^2 - 2|S(\omega)| \cos(\theta(\omega)), \quad (3.22)$$

where $\theta(\omega)$ is the phase angle of $S(\omega)$ in the complex plane at frequency $\omega$.\(^7\) To focus on the choice of $S(\omega)$, I substitute this expression into the objective function and rewrite problem (3.15) as

$$\min_{|S(\omega)|, \cos(\theta(\omega))} \int_{-\pi}^{\pi} \left[ (1 + |S(\omega)|^2 - 2|S(\omega)| \cos(\theta(\omega))) f_n(\omega) + |S(\omega)|^2 f_{xnc}(\omega) \right] d\omega, \quad (3.23)$$

subject to (3.10). Thus, the policymaker’s problem is reformulated as choosing function $|S(\omega)|$ and $\cos(\theta(\omega))$ in the frequency domain separately to minimize the loss function. This is intuitive because complex $S(\omega)$ is uniquely identified by its modulus and phase angle.

Notice that Bode’s constraint is a restriction on the modulus of the sensitivity function. In equation (3.23), modulus $|S(\omega)|$ is subject to (3.10), but phase angle $\theta(\omega)$ or function $\cos(\theta(\omega))$ is not constrained. Since both $|S(\omega)|$ and $f_n(\omega)$ are nonnegative at all frequencies, the loss function (3.23) is nonincreasing in $\cos(\theta(\omega))$.

\(^7\)Phase angle refers to the angular component of the polar coordinate representation of complex number.
Therefore, the optimal solution is to set $\cos(\theta(\omega)) = 1$, $\forall \omega \in [-\pi, \pi]$. It then follows that for the optimal policy rule with nonfiltered data,

$$|T(\omega)|^2 = (1 - |S(\omega)|)^2. \quad (3.24)$$

Given $\cos(\theta(\omega)) = 1$, $\theta(\omega) = 0$, $\forall \omega \in [-\pi, \pi]$. Thus, the sensitivity function $S^E(\omega)$ associated with the optimal policy rule turns out to be a real function. This is different from the noise-free benchmark case in which $S^B(\omega)$ is not necessarily real. A sensitivity function characterizes how the feedback control transforms $x^{nc}_t$ frequency by frequency to obtain $x^c_t$. In the frequency domain, "transform" means both gain and phase shifts at each frequency. A real sensitivity function implies no phase shifts or that all phase shifts are canceled out in the transform. The reason for $S^E(\omega)$ to be real is because $T^E(\omega)$ matters in this noisy control problem. As one can see from (3.15), a good control should have both $|S(\omega)|$ and $|T(\omega)|$ at all frequencies as small as possible. If $S(\omega)$ is not real and hence includes indelible phase shifts to $x^{nc}_t$, it will also induce indelible phase shifts to $n_t$. This will lead to a larger $|T(\omega)|$ than in the absence of such shifts. Intuitively, phase shifts also cause side noise effects, and therefore should prevented when possible.\(^9\)

\(^8\)In principle, if $f_n(\omega)$ vanishes to zero at some frequencies, then there may exist other solutions that $\cos(\theta(\omega)) \neq 1$ at these frequencies. This is not the generic case, so I withdraw from this technical issue.

\(^9\)Section B.2 of the Appendix provides a further example to illustrate the intuition for the real sensitivity function when accounting for measurement error.
Given equation (3.24), the policymaker’s problem (3.23) reduces to

\[
\min_{|S(\omega)|} \int_{-\pi}^{\pi} \left[ (1 - |S(\omega)|)^2 f_n(\omega) + |S(\omega)|^2 f_{xnc}(\omega) \right] d\omega, \quad (3.25)
\]

subject to Bode’s constraint (3.10). Measurement error stochastically disturbs the performance of any chosen policy rule in the time domain. However, the policy decision problem can be represented as a deterministically perturbed optimization problem in the frequency domain as (3.25): equation (3.25) is the same as benchmark (3.11) except that it includes the first term \((1 - |S(\omega)|)^2 f_n(\omega)\), which is deterministic and captures frequency-specific costs of the control. The effects of stochastic measurement error are now characterized by a deterministic control cost function. This representation greatly simplifies calculation in the frequency domain.

To solve problem (3.25), let \(\lambda^E\) be the Lagrangian multiplier associated with Bode’s constraint. The optimal solution gives the spectral density function of the controlled state variable as

\[
f_{x^E}(\omega) = \lambda^E + (1 - |S^E(\omega)|) f_n(\omega), \quad (3.26)
\]

with the Lagrangian multiplier

\[
\lambda^E = (|S^E(\omega)| - 1) |S^E(\omega)| f_n(\omega) + |S^E(\omega)|^2 f_{xnc}(\omega). \quad (3.27)
\]

Equation (3.27) is also the first order conditions for the optimization problem (3.25).
Therefore, the controlled spectral density $f_{x^E}(\omega)$ contains one constant term $\lambda^E$ and the other component that varies across frequencies. This can be interpreted as follows. First, the optimal policy rule still tries to flatten the uncontrolled spectral density by targeting the constant function $\lambda^E$. Unlike the noise-free or naive-policy target \(3.19\) which is solely determined by system disturbance attenuation, the spectral target $\lambda^E$ comprehensively accounts for the tradeoffs between stabilizing control effects and side noise effects, as shown in \(3.27\). Second, even though, the optimal control comes at an additional cost, captured by the second term of \(3.26\). For example, when the control shifts the uncontrolled density $f_{x^nc}(\omega)$ down at some frequency $\omega$, i.e. $|S^E(\omega)| < 1$, the stronger the control is the more side noise $(1 - |S^E(\omega)|) f_n(\omega)$ it brings into the system. When it lifts $f_{x^nc}(\omega)$ up, i.e. $|S^E(\omega)| > 1$, part of the control effect is offset by a downward force due to the fact that $(1 - |S^E(\omega)|) f_n(\omega) < 0$ in this case. Only when the control exerts no impacts on $f_{x^nc}(\omega)$, i.e. $|S^E(\omega)| = 1$, this additional cost goes to zero. In other words, with noisy data, the effectiveness of the control is reduced in the sense that the control fails to reach its spectral target; at frequencies where the control is more forceful, the larger the deviations will be from the target.

To complete the characterization of the solution, I solve for $|S^E(\omega)|$ from \(3.27\),

$$
|S^E(\omega)| = \frac{f_n(\omega) + \sqrt{f_n(\omega)^2 + 4\lambda^E (f_{x^nc}(\omega) + f_n(\omega))}}{2 (f_{x^nc}(\omega) + f_n(\omega))}, \quad (3.28)
$$
and determine $\lambda^E$ by substituting (3.28) into Bode’s constraint (3.10),

$$\int_{-\pi}^{\pi} \ln \left[ \frac{f_n(\omega) + \sqrt{f_n(\omega)^2 + 4\lambda^E (f_{xnc}(\omega) + f_n(\omega))}}{2 (f_{xnc}(\omega) + f_n(\omega))} \right]^2 d\omega = K_s.$$  

The Lagrangian multiplier $\lambda^E$, which is also the spectral target, is important for understanding the optimal policy rule with nonfiltered data. Several observations follow. First, $\lambda^E > 0$. This is guaranteed by Bode’s constraint; if $\lambda^E < 0$, then (3.27) implies that $|S^E(\omega)| < 1$ at all frequencies, which contradicts the fact that $K_s \geq 0$. Second, as shown in the numerical exercises of Section 3.4, $\lambda^E \geq \lambda^N$ in general. The optimal policy rule targets a higher constant spectral density and hence is less aggressive than the naive policy rule. Third, in the special case where measurement error is absent, $f_n(\omega) = 0, \forall \omega$, formula (3.26) and (3.28) will give exactly the same controlled spectral density function and sensitivity function as obtained in the noise-free benchmark, i.e. $f_{xE}(\omega)$ and $|S^E(\omega)|$ reduce to $f_{xB}(\omega)$ and $|S^B(\omega)|$.11

As characterized above, the optimal policy rule with nonfiltered data behaves differently from the naive policy rule in the frequency domain. A numerical comparison between them in a parameterized model will be presented in Section 3.4. But the main message has been clear here that there is a need to account for measurement error in the design of stabilization policies.

---

10Since $\lambda^E$ is positive, the other root of (3.27) is negative and therefore is not a reasonable solution.  
11Section B.3 of the Appendix shows that $f_{xE}(\omega)$ and $|S^E(\omega)|$ reduce to $f_{xB}(\omega)$ and $|S^B(\omega)|$ when there is no measurement error in the data.
Optimal Policy Rule with Filtered Data

Another natural method to address measurement error is to filter the data so that any chosen feedback rule can work more accurately. Suppose that the policymaker still sticks to the policy rule which is optimal for the benchmark noise-free system, but feeds back to the filtered data instead. I label this type of policies as optimal policy rule with filtered data.

Assume that the policymaker uses a linear filter $M(L)$,

$$\hat{x}_t = M(L)x_t^*.$$  \hfill (3.30)

He then applies the benchmark control $F^B(L)$ to the filtered data $\hat{x}_t$,

$$u_{t-1} = -F^B(L)M(L)x_{t-1}^*.$$  \hfill (3.31)

Therefore, the optimal policy rule with filtered data is

$$F^C(L) = F^B(L)M(L).$$  \hfill (3.32)

The associated sensitivity function is the Fourier transform of

$$S^C(L) = [(1 - M(L))(1 - A(L)L) + M(L)W(L)]^{-1}(1 - A(L)L),$$  \hfill (3.33)
which has

\[ |S^C(\omega)| = \left| \frac{1 - A(\omega)e^{-i\omega}}{(1 - M(\omega))(1 - A(\omega)e^{-i\omega}) + M(\omega)W(\omega)} \right|. \tag{3.34} \]

The spectral density function for the system under the control of \( F^C(L) \) is

\[ f_{x^C}(\omega) = \frac{\sigma_e^2|W(\omega)|^2 + \sigma_n^2|D(\omega)|^2|M(\omega)|^2|1 - W(\omega) - A(\omega)e^{-i\omega}|^2}{2\pi((1 - M(\omega))(1 - A(\omega)e^{-i\omega}) + M(\omega)W(\omega))^2}. \tag{3.35} \]

Applying the optimal policy rule with filtered data \( F^C(L) \), as shown in (3.35), the effects of the linear filter \( M(L) \) are twofold: it affects the way the benchmark control \( F^B(L) \) transforms the spectral density functions of the uncontrolled state variable \( x^nc_t \) and measurement error \( n_t \), and it also changes the way control noise enters into the system. The former is captured by \( M(\omega) \) in the denominator of (3.35), and the latter by the second term of (3.35).

It is clear that the filter \( M(L) \) plays the central role in this policy scenario. If the filter is set to \( M(L) = 1 \), i.e. no filtering, then the above results immediately reduce to be the same as those of the naive policy rule. For the rest of this subsection, I will discuss the choice of the optimal filter \( M(L) \). I argue that the Wiener filter is the most natural choice for the policymaker both because of its efficiency at filtering out measurement error and for its simple formula that is convenient for policy analysis in the frequency domain.

Since the data \( x^*_t \) is observed only up to time \( t \), the Wiener filter employed by the policymaker must be \textit{causal} in the sense that it is restricted to the one-sided
form,

\[ M(L) = \sum_{j=0}^{\infty} m_j L^j. \]  

(3.36)

The Wiener filter is efficient by the minimum mean-square error (MMSE) criterion, i.e., it is such that \( \text{MMSE} \equiv \mathbb{E} [(x_t - \hat{x}_t)^2] \) is minimized. The proof may be found in Priestley (1981, Chapter 10). Further, to give an explicit formula of the causal Wiener filter in the context of this chapter, suppose that the spectral density function of observational \( x^* \), \( f_{x^*}(\omega) = f_{xnc}(\omega) + f_n(\omega) \), satisfies the condition

\[ \int_{-\pi}^{\pi} \ln (f_{x^*}(\omega)) \, d\omega > -\infty. \]

It then admits a canonical factorization of the form

\[ f_{x^*}(\omega) = |H_{x^*}(e^{-i\omega})|^2, \]  

(3.37)

where \( H_{x^*}(e^{-i\omega}) \) is a "backward transform," a one-sided Fourier series involving only positive powers of \( \{e^{-i\omega}\} \). Thus, the \( z \)-polynomial \( H_{x^*}(z) = \sum_{j=0}^{\infty} h_j z^j \) has no zeros inside the unit circle. The causal Wiener filter is given by the Wiener-Kolmogorov formula\(^\text{13}\)

\[ M(\omega) = \left[ \frac{f_{xnc}(\omega)/H_{x^*}^*(\omega)}{H_{x^*}(\omega)} \right]_+ \]  

(3.38)

where \([ \cdot ]_+\) is the annihilation operator and superscript * again stands for the complex

---

\(^{12}\)If the Wiener filter is in the two-sided form, \( M(L) = \sum_{j=-\infty}^{\infty} m_j L^j \), then it is called non-causal; if defined with finite order of lag polynomial \( M(L) \), then called a finite impulse response (FIR) Wiener filter. Especially, the non-causal Wiener filter is such that in our case,

\[ M(\omega) = \frac{f_x(\omega)}{f_{x^*}(\omega)} = \frac{f_{xnc}(\omega)}{f_{xnc}(\omega) + f_n(\omega)} = \frac{\sigma_n^2 |W(\omega)|^2}{\sigma_e^2 |W(\omega)|^2 + \sigma_n^2 |D(\omega)|^2 |1 - A(\omega) e^{-i\omega}|^2}. \]

\(^{13}\)A equivalent representation in terms of \( z \)-polynomials is that \( M(z) = \frac{1}{H_{x^*}(z)} [f_{xnc}(z)/H_{x^*}^*(z)]_+ \).
Notice that the spectral density function of the filtered variable $\hat{x}$ is

$$f_{\hat{x}}(\omega) = |M(\omega)|^2 f_{x^*}(\omega).$$  \hspace{1cm} (3.39)

The Wiener filter places weights on the spectrum of $x_t^*$ frequency by frequency to adjust for the spectral density function of $\hat{x}$; $\hat{x}$ is the best predictor of the unobserved state variable $x_t$. It easy to verify that $|M(\omega)| \leq 1$ at all frequencies and that $|M(\omega)|$ is small at frequencies where the noise is strong. Intuitively, to approximate the spectral density function of the unobservable $x_t$, the filter pushes the spectral density of the observable $x_t^*$ down since the policymaker knows that measurement error $n_t$ contributes variance to $x_t^*$ at all frequencies; and, at frequencies where data noise is stronger the filter needs to remove more. However, even the optimal Wiener filter is not able to completely remove the noise. Thus, when the benchmark policy rule $F^B(L)$ is applied to the filtered data, it still introduces undesirable side noise into the system.

Optimal policy rule with nonfiltered data $S^E(\omega)$ and that with filtered data $S^C(\omega)$ represent two distinct ways of dealing with measurement error. One is to adjust the policy design, while the other is to process the data. Although it is hard to analytically characterize the differences of their performance in the current abstract model, numerical exercises developed in Section 3.4 will show their relative performance in the frequency domain.
3.3 Robust Policy Rule

Now let me introduce model uncertainty. In the current framework, there are two potential types of model uncertainty: uncertainty about structure of the state variable $x_t$ and uncertainty about that of measurement error $n_t$. Although models for unobservable $x_t$ and $n_t$ are based on the same information set $x_t^*$, the two types of model uncertainty are conceptually different. In this section, I will primarily consider model uncertainty with respect to the state variable $x_t$ because it represents a major limit in the policymaker’s knowledge about the structure of the economy. However, the analysis presented here can be immediately adopted to deal with uncertainty about the model of measurement error.

Specifically, the policymaker does not know the true model of $x_t$ but knows that it is close to a baseline model. Assume model uncertainty is the following regarding the spectral density function of the uncontrolled system,

$$
\int_{-\pi}^{\pi} \left[ f_{xnc}(\omega) - \tilde{f}_{xnc}(\omega) \right]^2 d\omega \leq \mu^2,
$$

(3.40)

where $f_{xnc}(\omega)$ is the unknown true model, $\tilde{f}_{xnc}(\omega)$ is the baseline model that the policymaker knows, and parameter $\mu > 0$ stands for the level of model uncertainty.\(^{14}\) This specification allows for various sources of model uncertainty such as uncertainty in parameter values, dynamic structure, and lag orders.

\(^{14}\)This specification is similar to Hansen and Sargent’s (2008, Chapter 8) formulation about the spectral density function of the innovations $W(L)\varepsilon_t$. If one is interested in uncertainty respect to measurement error model, an analogous specification would be

$$
\int_{-\pi}^{\pi} \left[ f_n(\omega) - \tilde{f}_n(\omega) \right]^2 d\omega \leq \mu^2,
$$
Following Hansen and Sargent (2008), robustness analysis is considered as a two-player zero-sum game. An adversarial agent is introduced, who chooses a model $f_{\text{nc}}(\omega, \mu)$ from the feasible neighborhood around the baseline $\bar{f}_{\text{nc}}(\omega)$ as specified by (3.40) with uncertainty level $\mu$ to maximize the loss function $\mathbb{E}[x_t^2]$, given the policymaker’s strategies. A primary agent, the policymaker, chooses a feedback control rule, represented by its sensitivity function $S(\omega, \mu)$, to stabilize the system given the potentially worst model chosen by the adversarial agent. In this sense, the robust policy rule is based on the worst-case analysis. However, playing against the adversarial agent the policymaker still uses the optimal control rule with nonfiltered data $S^E(\omega, \mu)$ to account for measurement error. Brock and Durlauf (2005) have shown that Nash and Stackelberg equilibria are approximately equivalent for the robustness games. I will hence use Nash equilibrium $(f_{\text{nc}}^R(\omega, \mu), S^E(\omega, \mu))$ as the solution concept, where $\mu$ is included in the equilibrium strategies to indicate the level of model uncertainty.

Assume that $\mu$ is sufficiently small, and therefore all feasible models are local. Following Brock and Durlauf’s (2005) approach, I will approximate the equilibrium $(f_{\text{nc}}^R(\omega, \mu), S^E(\omega, \mu))$ for the robust game around the baseline solution $(\bar{f}_{\text{nc}}(\omega), S^E(\omega, 0))$. This exercise is of particular interest in the frequency domain, since it allows us to explicitly characterize the policymaker’s frequency-specific marginal reactions to model uncertainty. Therefore, I will be able to explore how the presence of measurement error influences the policymaker’s reactions to model uncertainty across frequencies and changes the robustness properties of the optimal still with parameter $\mu$ measuring the degree of model uncertainty.
policy rule with nonfiltered data.

In the equilibrium, given the primary agent’s choice \( S^E(\omega, \mu) \), the adversarial agent solves

\[
\max_{f_{xnc}(\omega, \mu)} \int_{-\pi}^{\pi} \left[ |1 - S^E(\omega, \mu)|^2 f_n(\omega) + |S^E(\omega, \mu)|^2 f_{xnc}(\omega, \mu) \right] d\omega, \tag{3.41}
\]

subject to constraint (3.40). Let \( \gamma \) be the Lagrangian multiplier associated with the constraint. The first order conditions regarding the choice of \( f_{xnc}(\omega, \mu) \) are

\[
|S^E(\omega, \mu)|^2 = 2\gamma \left[ f^R_{xnc}(\omega, \mu) - \bar{f}_{xnc}(\omega) \right]. \tag{3.42}
\]

The objective function (3.41) is increasing in \( f_{xnc}(\omega, \mu) \) at each frequency. Thus, constraint (3.40) is binding in the solution and \( f^R_{xnc}(\omega, \mu) \geq \bar{f}_{xnc}(\omega) \). The decision of the adversarial agent in the Nash equilibrium then follows as

\[
f^R_{xnc}(\omega, \mu) = \bar{f}_{xnc}(\omega) + \mu r(\omega, \mu), \quad \text{with} \quad r(\omega, \mu) \equiv \frac{|S^E(\omega, \mu)|^2}{\sqrt{\int_{-\pi}^{\pi} |S^E(\omega, \mu)|^4 d\omega}}. \tag{3.43}
\]

Approximate \( f^R_{xnc}(\omega, \mu) \) around the baseline \( \bar{f}_{xnc}(\omega) \),

\[
f^R_{xnc}(\omega, \mu) = \bar{f}_{xnc}(\omega) + \mu r(\omega, 0) + o(\mu), \tag{3.44}
\]

with

\[
r(\omega, 0) \equiv \frac{|S^E(\omega, 0)|^2}{\sqrt{\int_{-\pi}^{\pi} |S^E(\omega, 0)|^4 d\omega}}, \tag{3.45}
\]
where \( S^E(\omega, 0) \) is just the optimal policy rule with nonfiltered data for the baseline model \( \tilde{f}_{xnc}(\omega) \) given by (3.28) and (3.29).

Notice that \( r(\omega, 0) \) is the adversarial agent’s marginal reaction to the level of model uncertainty \( \mu \) at frequency \( \omega \). It is related to how much control used by the policymaker at frequency \( \omega \) relative to the total amount used over the whole interval \([\pi, \pi]\) under the baseline model. The adversarial agent’s choice is to deviate more from the baseline model at frequencies where the policymaker uses control more. In this way, the adversarial agent disturbs the policymaker’s stabilization strategy as much as possible. From the policymaker’s perspective, \( f_{xnc}^R(\omega, \mu) \) specifies the worst situation he will face when model uncertainty is of level \( \mu \).

Now turn to the equilibrium strategy \( S^E(\omega, \mu) \) of the policymaker. Recall that the first order conditions for his optimization problem (3.25) under model \( f_{xnc}^R(\omega, \mu) \) are

\[
\left( |S^E(\omega, \mu)| - 1 \right) |S^E(\omega, \mu)| f_n(\omega) + |S^E(\omega, \mu)|^2 f_{xnc}^R(\omega, \mu) = \lambda^E(\mu). \tag{3.46}
\]

Differentiate both sides with respect to \( \mu \) to solve for \( \frac{\partial |S^E(\omega, \mu)|}{\partial \mu} \) and evaluate at \( \mu = 0 \):

\[
\frac{\partial |S^E(\omega, 0)|}{\partial \mu} = \frac{\frac{\partial \lambda^E(0)}{\partial \mu} - |S^E(\omega, 0)|^2 r(\omega, 0)}{(2|S^E(\omega, 0)| - 1) f_n(\omega) + 2|S^E(\omega, 0)|\tilde{f}_{xnc}(\omega)}, \tag{3.47}
\]

where I use the fact that \( f_{xnc}^R(\omega, 0) = \tilde{f}_{xnc}(\omega) \) and that \( \frac{\partial f_{xnc}^R(\omega, 0)}{\partial \mu} = r(\omega, 0) \) implied by (3.43). To pin down \( \frac{\partial |S^E(\omega, 0)|}{\partial \mu}, \frac{\partial \lambda^E(0)}{\partial \mu} \) in (3.47) has yet to be determined. Differentiate Bode’s constraint with respect to \( \mu \) and substitute \( \frac{\partial |S^E(\omega, \mu)|}{\partial \mu} \) with (3.47). Then, I can
solve for \( \frac{\partial \lambda^E(\mu)}{\partial \mu} \) and evaluate it at \( \mu = 0 \):

\[
\frac{\partial \lambda^E(0)}{\partial \mu} = \int_{-\pi}^{\pi} \frac{|S^E(\omega,0)|^2 r(\omega,0)}{|S^E(\omega,0)||2|S^E(\omega,0)|-1| f_n(\omega) + 2|S^E(\omega,0)||f_{xnc}(\omega)|} \, d\omega.
\] (3.48)

Therefore, in the Nash equilibrium the policymaker’s strategy can be characterized as

\[
|S^E(\omega, \mu)| = |S^E(\omega, 0)| + \mu \frac{\partial |S^E(\omega, 0)|}{\partial \mu} + o(\mu),
\] (3.49)

with \( \frac{\partial |S^E(\omega, 0)|}{\partial \mu} \) determined by (3.47) and (3.48).

The policymaker’s equilibrium choice \( S^E(\omega, \mu) \) is just the robust policy rule. Several observations follow. First, by linearization, the term \( \frac{\partial |S^E(\omega, 0)|}{\partial \mu} \) in (3.47) and (3.49) describes the policymaker’s marginal reaction by frequency to model uncertainty. This reaction at frequency \( \omega \) are driven by two forces. On one hand, including model uncertainty increases the constant density level that the optimal policy rule targets, and hence reduces control aggressiveness at all frequencies. This is captured by \( \frac{\partial \lambda^E(0)}{\partial \mu} \), which is positive and constant over \( \omega \). On the other hand, the robust policy rule \( S^E(\omega, \mu) \) also reacts to the adversarial agent’s adjustment on the uncontrolled spectral density. Since the adversarial agent tends to increase the uncontrolled spectral density by rate \( r(\omega, 0) \), the policymaker will change his control in the opposite direction by \( |S^E(\omega, 0)|^2 r(\omega, 0) \); the more the original control \( |S^E(\omega, 0)| \) is used, the more the marginal reaction should be. In the end, both parts of the reaction are adjusted by the denominator \( (2|S^E(\omega, 0)|-1) f_n(\omega) + 2|S^E(\omega, 0)||f_{xnc}(\omega)| \) to satisfy Bode’s constraint. Second, the analysis provided here
is based on the policymaker using the optimal policy rule with nonfiltered data, but it can be extended to the cases using the naive policy rule and the optimal policy rule with filtered data with some complications. This indicates that the way of dealing with measurement error matters for both optimal and robust policies. Finally, the characterization of the robust policy rule as (3.49) illustrates the interaction between model uncertainty and measurement error in the design of stabilization policy rules. This is absent in, for example, Brock and Durlauf (2005) and new to the literature. In equation (3.49), \( \frac{\partial |sE(\omega,0)|}{\partial \mu} \) is marginal reaction by frequency to model uncertainty of level \( \mu \). However, it is clear from (3.47) and (3.48) that this reaction rate depends on the behavior of measurement error \( f_n(\omega) \) in the frequency domain via several different channels. In this fashion, the interplay between the concerns over these two sources of uncertainty is characterized frequency by frequency. The numerical exercises in the next section will picture such effects more intuitively.

### 3.4 Some Applications: Monetary Policy Evaluation

In this section, I apply the general theory developed in the previous sections to evaluate monetary policy rules. First, I embed the conventional two-equation Keynesian monetary model into the scalar AR(1) framework so that I can work with the analytic devices constructed above directly. Then, based on the parameterized AR(1) model, I perform numerical experiments to assess several aspects of the measurement error effects on monetary policy design.

Typically, models employed in the literature of monetary policy evaluation
contain two structural equations. One specifies the Phillips curve relating the
output gap and inflation, and the other specifies the IS curve relating the real
interest rate to the output gap. A simple version of the well-known Rudebusch and
Svensson (1999) model which is widely used in previous studies can be written in
the form of a purely backwards-looking system:

\[ \pi_t = \rho \pi_{t-1} + by_{t-1} + \varepsilon_t \]  \hspace{1cm} (3.50)

\[ y_t = \theta y_{t-1} - \gamma (i_t - \pi_t) + e_t \] \hspace{1cm} (3.51)

where \( y_t \) stands for the gap between output and potential output, \( \pi_t \) is inflation
and \( i_t \) is nominal interest rate. A Taylor (1993) type rule sets interest rate as

\[ i_t = a_{\pi} \pi_t + a_y y_t, \] or more generally as

\[ i_t = a_{\pi}(L)\pi_t + a_y(L)y_t. \] \hspace{1cm} (3.52)

This linear feedback rule of interest rate (3.52), together with the IS curve (3.51),
implies the output gap \( y_t \) as

\[ y_t = -\beta(L)\pi_t + \phi(L)e_t, \] \hspace{1cm} (3.53)

with \( \beta(L) \) and \( \phi(L) \) determined during the rearrangement. Then, the economic
environment can be considered as a one-state-variable-one-control system, with
the system specified by the Phillips curve (3.50) and the control by (3.53).

In this essay, however, I still need to tailor the model slightly to have the inter-
pretation of monetary policymaking as a feedback control problem in the presence of noisy information; I do this by restricting the control so that

\[ y_{t-1} = -\beta(L)\pi^*_{t-1} \]  
(3.54)

\[ \pi^*_t = \pi_t + n_t \]  
(3.55)

where the error term \( n_t \equiv -\frac{\phi(L)e_t}{\beta(L)} \) can be simply understood as a process of measurement error in this context. To simplify the analysis, assume that \( \rho < 1 \) in (3.50) so that the system is stationary and that \( \beta(L) \) is a one-sided lag polynomial. Further, let measurement error \( n_t \) be white noise with variance \( \sigma_n^2 \), in light of Orphanides’ (2003) evidence that the inflation noise can be adequately modeled as a serially uncorrelated process. Fundamental innovations \( \varepsilon_t \) have variance \( \sigma^2_\varepsilon \). Then, ratio \( \nu \equiv \frac{\sigma_n^2}{\sigma^2_\varepsilon} \) represents the relative strength of measurement error.

To focus on the performance of policy rules in inflation stabilization, consider the loss function

\[ V(\pi) = \mathbb{E}[\pi_t^2]. \]  
(3.56)

This preference means that the policymaker is, in King’s (1997) words, an "inflation nutter," who does not care about output stabilization. It can also be justified as inflation targeting in practice. The policy decision is then to design a feedback rule \( \beta(L) \) of (3.54) in the presence of measurement error \( n_t \) to minimize \( V(\pi) \), given the AR(1) economic model (3.50). And I will focus on the study of \( \beta(L) \) and its sensitivity function in the frequency domain.
Admittedly, the model outlined here is highly stylized. Abstract from many practical issues, it highlights the importance to account for measurement error properly in monetary policy evaluation. Measurement error \( n_t \) can be broadly interpreted as an inaccurate control in this framework, which may have emerged in various ways. For example, \( n_t \) may arise directly from the mismeasurements of aggregate price levels and price changes, but can also result from misspecifications of the IS curve due to the lack of exact knowledge about monetary transmission mechanism, as implied by the appearance of IS error \( e_t \). The numerical exercises below convey the idea how to quantitatively and intuitively assess the frequency-specific effects of this inaccurate control.

**Policy Rules without Lagged Terms**

A natural starting point is to consider the case in which policy rules depend only on the current observation of the state variable; lagged terms are excluded. This is interesting because the effects of measurement error can be characterized by one single coefficient parameter in this simple case. Also, simple instrument rules without lagged terms [e.g., Taylor (1993)] or with only a few lags [e.g., Onatski and Williams (2003)] have received much attention in the related literature. The results developed below are not trivial but rather relevant to policy practice.

Ignoring all pervious information, control (3.54) takes the form of

\[
y_t = -\beta \pi_{t-1}^*.
\]  

(3.57)
The solution to this case can be worked out explicitly. If the policymaker is not aware of measurement error $n_t$, he will use the naive policy rule $y_{t-1} = -\beta^N \pi^*_{t-1}$ with

$$\beta^N = \frac{\rho}{b}. \quad (3.58)$$

Otherwise, integrating his knowledge of $n_t$ into policy decision, the policymaker’s optimal policy rule with nonfiltered data in this case turns out to be $y_{t-1} = -\beta^E \pi^*_{t-1}$ with

$$\beta^E = -\frac{1}{2} \frac{\rho b \sigma_n^2}{\sigma_n^2 + \sigma^2_{\epsilon}} \left[ (1 - \rho^2) \sigma_n^2 + \sigma^2_{\epsilon} - \sqrt{[(1 - \rho^2) \sigma_n^2 + \sigma^2_{\epsilon}]^2 + 4 \rho^2 \sigma_n^2 \sigma^2_{\epsilon}} \right]. \quad (3.59)$$

One important observation can be made here. The optimal policy rule $\beta^E$ adjusted for the measurement turns out to be less responsive than the naive policy rule $\beta^N$ that assumes measurement is exact. This is because

$$|\beta^E| \leq |\beta^N|. \quad (3.60)$$

Details of the derivation and comparison of these policy rules can be found in Section B.4 of the Appendix. But the implication of this result is straightforward; if the policymaker is not confident in the accuracy of his data, excess activeness is harmful.

Figure 3.1 draws the naive policy rule $\beta^N$ and optimal policy rule $\beta^E$ when $v$ varies from 0 to 10 and $\rho$ from 0 to 1. Here, $b$ is calibrated to be 0.1. Consistent with intuition, Figure 3.1 shows that whether to account for measurement error or not is
Figure 3.1: Naive Policy Rules and Optimal Policy Rules without Lagged Terms

Note: \( v \equiv \frac{\sigma_n^2}{\sigma^2} \) represents the relative strength of measurement noise, \( \rho \) is the parameter of AR(1) Phillips curve, and \( b = 0.1 \). \( \beta^N \) is the response of the naive policy rule to the current observation and \( \beta^E \) is that of the optimal policy rule.

irrelevant, \( \beta^N = \beta^E \), only when measurement error is trivially absent, \( \sigma_n^2 = 0 \) (i.e., \( v = 0 \)), or when the controlled system is a white noise process itself and hence no control should be used anyway, \( \rho = 0 \). Otherwise, the more serious measurement error is (i.e., the larger \( v \) is) or the more persistent the uncontrolled system is (i.e., the larger \( \rho \) is), the less aggressive the optimal policy rule \( \beta^E \) will be when compared with the naive rule \( \beta^N \). The role of noise level \( v \) is straightforward; when control induces side noise, the policymaker will use control more cautiously when he is aware of the stronger side effects. However, the role of system persistence \( \rho \) needs some explanation: A persistent uncontrolled system requires a very active feedback to the current observation so that the volatility inherited from previous periods can
be canceled out; in other words, the stabilizing control effects of any policy action are weak relative to side noise effects, and thus, the policy works better by using a less responsive feedback rule to restrain the side noise effects.

In addition, in Figure 3.1 the optimal policy rule $\beta^E$ is very steep when $\rho$ is high and $v$ is small. In persistent model, even a small level of measurement noise may cause a large change in the optimal policy rule design. This illustrates the interaction between model uncertainty and measurement noise in affecting optimal policymaking. Here, model uncertainty is about the different values of $\rho$. Under different models, the policymaker adjusts policy aggressiveness for measurement error in rather different degrees. I will explore this point more generally in the frequency domain later.

It is also interesting to investigate the spectral properties of the naive rule $\beta^N$ and the optimal rule $\beta^E$ in this non-lagged-term case. The sensitivity function associated with $\beta^N$ is

$$|S^N(\omega)| = \sqrt{1 - 2\rho \cos(\omega) + \rho^2}, \quad (3.61)$$

and the sensitivity function associated with $\beta^E$ is

$$|S^E(\omega)| = \frac{\sqrt{1 - 2\rho \cos(\omega) + \rho^2}}{\sqrt{1 + 2(b\beta^E - \rho) \cos(\omega) + (b\beta^E - \rho)^2}}, \quad (3.62)$$

At frequencies where $\cos(\omega) > \frac{1}{2}(\rho - b\beta^E)$, $|S^E(\omega)| \geq |S^N(\omega)|$, while at frequencies where $\cos(\omega) < \frac{1}{2}(\rho - b\beta^E)$, $|S^E(\omega)| < |S^N(\omega)|$, when $\rho > b\beta^E$; and vice versa when
bβ^E \geq \rho$. The optimal rule is less sensitive than the naive rule, $|\beta^E| \leq |\beta^N|$, in general as shown above, but over a range of frequencies the former can actually be more aggressive than the latter. It is right in this sense that spectral analysis is informative when one is interested in how measurement error changes the allocation of control effects across frequencies.

**Optimal Policy Rule with Nonfiltered Data**

In the general case without the restriction on the order of lag polynomial, a monetary policy rule has sensitivity function and complementary sensitivity function as

$$S(\omega) = \frac{1 - \rho e^{-i\omega}}{1 - (\rho - b\beta(\omega))e^{-i\omega}} \quad \text{and} \quad T(\omega) = \frac{b\beta(\omega)e^{-i\omega}}{1 - (\rho - b\beta(\omega))e^{-i\omega}}. \quad (3.63)$$

The measurement error process $n_t$ and uncontrolled inflation $\pi^{nc}_t$ have spectral density functions as

$$f_n(\omega) = \frac{\sigma_n^2}{2\pi}, \quad \text{and} \quad f_{\pi^{nc}}(\omega) = \frac{\sigma_\pi^2}{2\pi(1 - 2\rho \cos(\omega) + \rho^2)}. \quad (3.64)$$

As established in Section 3.2, the optimal policy rule with nonfiltered data is then pinned down by its sensitivity function

$$|S^E(\omega)| = \frac{\sigma_n^2}{2\pi} + \sqrt{\left(\frac{\sigma_n^2}{2\pi}\right)^2 + 4\lambda^E \left(\frac{\sigma_n^2}{2\pi} + \frac{\sigma_\pi^2}{2\pi(1 - 2\rho \cos(\omega) + \rho^2)}\right)} \quad (3.65)$$
where the constant spectral target $\lambda^E$ is determined by numerically solving from condition (3.29), i.e. $\int_{-\pi}^{\pi} \ln[|S^E(\omega)|^2] d\omega = 0$, with $f_n(\omega)$ and $f_{\pi nc}(\omega)$ given above.

I employ the following parameterizations. Following Brock et al (2007), I calibrate $\rho = 0.9$ and $b = 0.1$. I also specify different levels of noise strength as: $\nu = 0$, measurement error is absent; $\nu = 0.4402$, measurement error is modest and it is derived from Orphanides’ (2003) estimates $\sigma_\varepsilon = 1.04$ and $\sigma_n = 0.69$; $\nu = 1$, measurement error has the same standard deviation as fundamental innovations $\varepsilon_t$; and $\nu = 4$, measurement error is serious and with the standard deviation twice as that of $\varepsilon_t$.

Figure 3.2 presents the sensitivity functions $S^E(\omega)$ associated with the optimal policy rules with nonfiltered data when exposed to the different levels of data
noise. The central message is that policy inertia is increasing in noise strength. To flatten spectral density, an optimal policy rule is to push the uncontrolled $f_{\pi^{nc}}(\omega)$ down (i.e., $|S^E(\omega)| < 1$) at low frequencies and to lift it up (i.e., $|S^E(\omega)| > 1$) at high frequencies in this AR(1) framework. When measurement error is present in the data, the policymaker has to make tradeoffs between stabilizing control effects and side noise effects in applying feedback control rules as explained in Section 3.2. In Figure 3.2, the optimal policy rule pushes $f_{\pi^{nc}}(\omega)$ down less at low frequencies and also lifts it up less at high frequencies as measurement error becomes stronger. Therefore, the total effect of measurement error is to make the optimal monetary policy rule less aggressive.

However, the policymaker does not reduce aggressiveness equally across frequencies. In Figure 3.2, the sensitivity function moves close to constant level 1 much more quickly at high frequencies than at low frequencies as the strength of measurement error strength increases. This essentially means that measurement error induces more cautiousness in the use of control at high frequencies than at low frequencies. Especially over a small range of medium frequencies, measurement error actually leads to more active policy control. The intuition behind as follows. In this AR(1) framework, the variance of the uncontrolled state variable concentrates at low frequencies, where system volatility is so high that the policymaker has to use control somewhat forcefully even if it brings some side effects. Even when measurement error is strong, these side noise effects are still small relative to the system volatility at these frequencies. In contrast, side noise effects quickly dominate stabilizing control effects at high frequencies, so the policymaker must
be much more cautious as data noise increases. Finally, at medium frequencies, the policymaker may have to be more aggressive because of the design limits – when the spectral density is pushed down at low frequencies less than the amount lifted up at high frequencies due to the unbalanced frequency-specific reactions, the density function has to popup at some medium frequencies to meet Bode’s constraint.

Figure 3.3 shows the spectral density functions of inflation $\pi_t^E$ under the control of the optimal policy rules at the different levels of measurement error. Without data noise the optimal control completely flattens the density function, as expected. In the presence of measurement error, the weaker measurement error is, the flatter the
controlled density function will be. The total variance of the controlled state variable, i.e. the area under the densities, is increasing in the strength of measurement error. Therefore, the optimal policy rule is more effective when data noise is lower. Across frequencies, the "waterbed effects" – the reduction of variance at some frequencies results in increases of variance at other frequencies – implied by Bode’s constraint are also evident from the Figure. On one hand, when measurement error is stronger, the controlled density function has a higher peak at low frequencies. This comes in two ways. First, the control is less aggressive so that the uncontrolled peak is pushed down less. Second, it brings stronger side noise effects when data noise is stronger. On the other hand, the optimal policy rule performs well at high frequencies in the sense that the controlled density is low and flat, and this does not change much when measurement error increases.

Figure 3.4 presents the assessment of aggressiveness of the optimal policy rule over different levels of measurement error or model persistence. Both notions of the AG measure and spectral target $\lambda^E$ are used. Given model persistence $\rho = 0.9$, panel (a) shows that the spectral target $\lambda^E$ is increasing in the strength of measurement error $v$, and panel (b) shows that the AG measure is decreasing in $v$. Both mean that when the measurement is noisier the optimal policy rule is less aggressive. It is also worth noting that both $\lambda^E$ and AG change at a decreasing rate over $v$. Panel (c) displays the spectral target $\lambda^E$ over varying model persistence $\rho$ in the cases with and without measurement error. In the absence of measurement error, the optimal policy rule always targets a constant spectral level. When modest measurement error is present, the target $\lambda^E$ increases in $\rho$. Thus, the target gap
between the two cases is also increasing. Intuitively, when model persistence $\rho$ increases, the stabilizing control effect of any given policy rule becomes weaker, and hence the policymaker has to increase its spectral target when measurement error is present and can cause side noise. In panel (d), even without measurement error the $AG$ aggressiveness is still increasing $\rho$. This is because the uncontrolled density has a more precipitous peak with a larger $\rho$. To target the same low spectral

Note: $AG$ is defined in $L^2$ norm $AG = \left[ \int_{-\pi}^{\pi} \left| 1 - |S^E(\omega)|^2 \right|^2 d\omega \right]^{\frac{1}{2}}$ and $\lambda^E$ is the constant spectrum target.
level essentially means a more aggressive control when the model becomes more persistent. Panel (d) also shows that the AG aggressiveness is always smaller when measurement error is present than when it is absent; again the gap is increasing in model persistence $\rho$. Thus, the optimal policy rule reacts to the same level of measurement error more in reducing aggressiveness when facing a more persistent model. Interestingly, AG is almost linear in $\rho$ for both cases.

**Optimal Policy Rule with Filtered Data**

For the optimal policy rule with filtered data, the most important step is to filter the data properly. Within the AR(1) framework, I first work out the Wiener filter explicitly, following Priestley (1981, Chapter 10), and then use it for numerical implementation.

Given that the unobservable $\pi_t$ is an AR(1) process and that measurement error $n_t$ is white noise, the spectral density function of observed $\pi_t^*$ is

$$f_{\pi^*}(\omega) = \frac{\sigma_z^2}{2\pi|1 - \rho e^{-i\omega}|^2} + \frac{\sigma_n^2}{2\pi}. \quad (3.66)$$

Its canonical factorization gives

$$f_{\pi^*}(z) = \frac{k(1 - \phi z)(1 - \phi z^{-1})}{2\pi(1 - \rho z)(1 - \rho z^{-1})}, \quad (3.67)$$

where $k = \rho \sigma_n^2 / \phi$ and $\phi$ is the root inside the unit circle of $z$-polynomial $f_{\pi^*}(z)$. In
this case,

\[ H_{\pi^*}(z) = \sqrt{\frac{k}{2\pi}} \left( \frac{1 - \phi z}{1 - \rho z} \right). \]  \hspace{1cm} (3.68)

As shown in Section B.5 of the Appendix, this implies that

\[ \left[ \begin{array}{c} f_{\pi}(z) \\ H_{\pi^*}(z) \end{array} \right] = \frac{\sigma^2_{\epsilon}}{(1 - \rho \phi)(1 - \rho z)\sqrt{2\pi k}}. \]  \hspace{1cm} (3.69)

Therefore, the causal Wiener filter (3.38) in the frequency domain can be written as

\[ |M(\omega)|^2 = \left[ \frac{\phi \sigma^2_{\epsilon}}{(1 - \rho \phi)\rho \sigma^2_{\pi}} \right]^2 \frac{1}{1 + \phi^2 - 2\phi \cos(\omega)}. \]  \hspace{1cm} (3.70)

Recall that the spectral density of the filtered variable \( \hat{\pi} \) is simply

\[ f_{\hat{\pi}}(\omega) = |M(\omega)|^2 f_{\pi^*}(\omega), \]  \hspace{1cm} (3.71)

so \( |M(\omega)|^2 \) are just weights placed on \( f_{\pi^*}(\omega) \) across frequencies to adjust for \( f_{\pi}(\omega) \).

Panel (a) in Figure 3.5 presents the adjustment weights placed by the Wiener filter on \( f_{\pi^*}(\omega) \). First, the Wiener weights are lower at all frequencies with weak data noise than with strong noise. For example, when measurement error is almost absent, \( \nu \to 0 \), the weights are almost equal to 1 everywhere, whereas the case of \( \nu = 4 \) has the lowest weighting curve. Second, the weights are high at low frequencies and low at high frequencies. This is because the AR(1) unobservable \( \pi_t \) amounts to a high fraction at low frequencies and a low fraction at high frequencies of volatility of the observable \( \pi^*_t \) when measurement error is white noise. Third, the
Note: $v \to 0$ is approximated by setting $v = 0.0001$. It captures the case when noise is almost absent.

weighting curve is not shifted down by the equal distance across frequencies when measurement error becomes stronger. For example, when $v$ goes from 0.4402 to 1, the weighting line moves more down at high frequencies than at low frequencies, because measurement error has stronger effects at high frequencies than at low frequencies. Panel (b) shows the filtered results at frequencies $[\pi/2, \pi]$. The spectral density function of the filtered $\hat{\pi}_t$ is very close to that of the unobservable $\pi_t$, so the Wiener filter works very well. After the policymaker applies the optimal policy rule
for the noise-free model to the filtered data, the associated sensitivity functions and policy performance are then shown in panels (c) and (d) of Figure 3.5. One can see from panel (c) that measurement error reduces the aggressiveness of policy control in general, and reduces it more at high frequencies than at low frequencies. Panel (d) indicates that the extent of the side noise effects brought into the system by the optimal policy rules with filtered data still increases in the strength measurement error.

**Robust Policy Rule**

To study the robustness properties of monetary policy rules, I need to specify the level of model uncertainty. Since this section mainly serves to communicate the method to track the interplay of measurement error and model uncertainty in policymaking, I would withdraw from the involvement in the data-based calibration of $\mu$ but simply set $\mu = 1$.\(^\text{15}\) I assume that the true model is local to the AR(1) baseline model as studied above. Thus,

$$f_{\pi \text{nc}}(\omega) = \frac{1}{2\pi(1 - 2\rho \cos(\omega) + \rho^2)}$$

and

$$\int_{-\pi}^{\pi} [f_{\pi \text{nc}}(\omega) - \bar{f}_{\pi \text{nc}}(\omega)]^2 d\omega \leq 1. \quad (3.72)$$

I maintain the assumption that measurement error is white noise with the different levels of strength $v$.

In Figure 3.6, panel (a) shows the adversarial agent’s marginal reactions $r(\omega, 0)$ across frequencies to model uncertainty. The adversarial agent knows that the

\(^{15}\text{For more rigorous empirical implementation, see Hansen and Sargent (2001, 2008) for details on how the level of model uncertainty } \mu \text{ can be calibrated quantitatively from historical data.}$$
policymaker will push the uncontrolled density down very forcefully and hence it is less effective to disturb at low frequencies. In contrast, at high frequencies the policymaker allows the uncontrolled density to move up, and then the adversarial agent’s disturbing actions will not be canceled out. Therefore, the adversarial agent puts most disturbing power at high frequencies. This renders the marginal reaction curves a similar shape as the sensitivity functions associated with the optimal rules shown in Figure 3.2.

Panel (b) shows how the policymaker will react. Consider the noise-free case first. There is only model uncertainty, and the policymaker knows that most uncertainty comes from high frequencies. Without much influence of the adversarial agent, the policymaker will not change his strategies very much at low frequencies. However, he will greatly reduce control used at high frequencies. To flatten the density function, the policymaker needs the uncontrolled AR(1) density to move up at high frequencies in the absence of model uncertainty. But, now the adversarial agent also lifts up the spectral density at high frequencies which helps the policymaker to reduce the control used. The consequence is that more control is moved to medium frequencies to meet Bode’s constraint. This explains the "m" shape of the policymaker’s marginal reaction curve.

In the presence of measurement error, the "m" shape is less obvious, although still present. A flat "m" shape marginal reaction curve around zero means that the policymaker only mildly reacts to model uncertainty at all frequencies. This is especially true for the case with strong data noise, $\nu = 4$. At low frequencies, it is not necessary to react very much as the adversarial agent puts little uncertainty here
anyway. Knowing that the policymaker tends to balance control used at medium and high frequencies by making the sensitivity function flatly close to constant 1 (see Figure 3.2) when measurement error is present, the adversarial agent also balances the allocation of model uncertainty at these frequencies as illustrated by panel (a) of Figure 3.6. The policymaker then has to reduce his reaction at both medium and high frequencies. Put differently, reacting actively to model uncertainty will increase control used at medium frequencies as shown in the noise-free case, but control comes with side noise effects when measurement error is strong. Thus, the policymaker simply chooses not to react that much. Reducing reaction at medium frequencies also leads to the reductions at high frequencies, according to Bode’s constraint. The whole point is that the presence of measurement error reduces the policymaker’s reaction to model uncertainty, and this is especially clear at medium and high frequencies.

Taking model uncertainty into consideration, the sensitivity functions associated with the robust policy rules are displayed in panel (d). Whether measurement error is present or not, the robust policy rules behave rather differently, especially at medium and high frequencies. The interaction between model uncertainty and measurement error leads to the performance of the robust policy rules as shown in panel (c).

In the light of Brainard’s (1967) argument that model uncertainty justifies cautious policy, I may conclude based on these exercises: (1) cautiousness resulting from model uncertainty is not equal across frequencies – the policymaker reacts to model uncertainty less at low frequencies than at others, and the reaction at
medium frequencies actually leads to active control in the contrast; (2) measurement error reduces the impacts of model uncertainty on policymaking – this is especially important at medium frequencies, which may overlap business cycle frequencies that monetary authorities potentially care most, in the sense that model uncertainty does not overthrow the general insights on monetary policymaking. Rather, model uncertainty should be assessed along with other forms of uncertainty such as data noise. In the environment of various uncertainty sources, the effects of model uncertainty may be less significant than in the situation focusing on model
Table 1: Comparison of Policy Scenarios under Different Levels of Measurement Error

<table>
<thead>
<tr>
<th>Level of Data Noise</th>
<th>Naive Policy</th>
<th>Optimal Policy, Nonfiltered</th>
<th>Optimal Policy, Filtered</th>
<th>Robust Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement Error : $v = 0$ (No Noise)</td>
<td>density level target $\lambda$</td>
<td>0.1592</td>
<td>0.1592</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>aggressiveness $AG$</td>
<td>1.5411</td>
<td>1.5411</td>
<td>1.5407</td>
</tr>
<tr>
<td></td>
<td>variance under control $\sigma^2_c$</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9999</td>
</tr>
<tr>
<td></td>
<td>high frequency variance contribution $\left[\pi/2, \pi\right]$</td>
<td>0.2500</td>
<td>0.2500</td>
<td>0.2500</td>
</tr>
<tr>
<td>Measurement Error : $v = 0.4402$ (Orphanides Estimate)</td>
<td>density level target $\lambda$</td>
<td>0.1592</td>
<td>0.1919</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>aggressiveness $AG$</td>
<td>1.5411</td>
<td>1.2608</td>
<td>1.0038</td>
</tr>
<tr>
<td></td>
<td>variance under control $\sigma^2_c$</td>
<td>1.1664</td>
<td>1.1336</td>
<td>0.9315</td>
</tr>
<tr>
<td></td>
<td>high frequency variance contribution $\left[\pi/2, \pi\right]$</td>
<td>0.3087</td>
<td>0.2411</td>
<td>0.1724</td>
</tr>
<tr>
<td>Measurement Error : $v = 1$</td>
<td>density level target $\lambda$</td>
<td>0.1592</td>
<td>0.2212</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>aggressiveness $AG$</td>
<td>1.5411</td>
<td>1.0981</td>
<td>0.8956</td>
</tr>
<tr>
<td></td>
<td>variance under control $\sigma^2_c$</td>
<td>1.3780</td>
<td>1.2560</td>
<td>0.9917</td>
</tr>
<tr>
<td></td>
<td>high frequency variance contribution $\left[\pi/2, \pi\right]$</td>
<td>0.3834</td>
<td>0.2346</td>
<td>0.1500</td>
</tr>
<tr>
<td>Measurement Error : $v = 4$</td>
<td>density level target $\lambda$</td>
<td>0.1592</td>
<td>0.2597</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>aggressiveness $AG$</td>
<td>1.5411</td>
<td>0.9464</td>
<td>0.8357</td>
</tr>
<tr>
<td></td>
<td>variance under control $\sigma^2_c$</td>
<td>1.7560</td>
<td>1.4199</td>
<td>1.1226</td>
</tr>
<tr>
<td></td>
<td>high frequency variance contribution $\left[\pi/2, \pi\right]$</td>
<td>0.5168</td>
<td>0.2274</td>
<td>0.1352</td>
</tr>
</tbody>
</table>

Note: In the case of $v = 0$, optimal control rule is calculated by setting $v = 0.0001$ to use the Wiener filter.
Comparing Policy Scenarios

As a final exercise, I compare the performance of naive policy rule, optimal policy rules with nonfiltered and filtered data, and robust policy rule in the current parameterized model. Recall that, in this setting, the naive policy rule is associated with the sensitivity function

\[ |S^N(\omega)| = \sqrt{1 - 2\rho \cos(\omega) + \rho^2}, \quad (3.73) \]

and has the spectral density function under control as

\[ f_{\pi N}(\omega) = \frac{\sigma^2}{2\pi} + \frac{\sigma^2_n}{2\pi} (1 - S^N(\omega))^2, \quad (3.74) \]

along with a constant level of spectral target \( \lambda^N = \frac{\sigma^2}{2\pi} \). The other three policy scenarios have already been examined as above.

Table 1 reports the comparison results under the different levels of measurement error. First, the optimal policy rules with nonfiltered and filtered data are less aggressive and have better performance than the naive policy rules. The conclusion on aggressiveness holds for both spectral target \( \lambda \) and AG measure under different levels \( \nu \) of measurement error. Meanwhile, the optimal policy rules reduce the total variances to lower levels than what the naive policy rules can reach. This indicates that failing to recognize measurement noise will lead to serious distortions in policy
performance. Second, the optimal policy rules with filtered data outperform those with nonfiltered data. The former give lower variances than the latter over the whole frequency domain as well as over the high frequency range, while the former are less aggressive than the latter in the overall AG measure. This may be interpreted as the power of the Wiener filter in filtering out measurement noise. The optimal policy rules with nonfiltered data introduce more side noise because of more aggressive control. Third, the robust policy rules are the least aggressive in both target $\lambda$ and measure AG, as expected. Interpreting these results, it is important to note that the robust policy rules are based on the least favorable model, which is different from the baseline model that the other rules are dealing with.

Figures 3.7 and 3.8 show the comparisons of the sensitivity functions and their
performance across the different policy scenarios when measurement error is modest, $v = 0.4402$. With regard to the sensitivity functions, the four policy rules place similarly strong control at low frequencies, although the robust policy is a little weaker. The main differences come from medium and high frequencies. For example, the robust policy rule is the least responsive at high frequencies and the most active at medium frequencies. The sensitivity functions of the optimal policy rules have very similar shapes except that the rule with nonfiltered data has its sensitivity function below that of the rule with filtered data. This leads to the result in Figure 3.8 that the optimal policy rule with filtered data outperforms the optimal rule with nonfiltered data. In Figure 3.8, the controlled spectral density function under the robust policy is the highest since it assumes the highest uncontrolled
spectral density rather than the AR(1) baseline density. Interestingly, the naive
policy rule does not work well in total-variance reduction, but at least it is effective
at low frequencies. Its inefficiencies arise at high frequencies. This again high-
lights the importance of studying frequency-specific effects of measurement error
in monetary policy evaluation.

3.5 Concluding Remarks

This chapter provides a framework for investigating how alternative feedback policy
rules behave in the presence of measurement errors and with respect to frequency-
specific performance. I argue the importance to recognize potential data noise
in policy decisionmaking, and show various ways to integrate this information
into the assessment of policies’ efficiency and robustness. Applied to monetary
policy evaluation, the numerical exercises draw insights on the frequency-specific
adjustments in the design of policy rules to the measurement error.

Admittedly, this essay contains weak points, which I would like to explore in
future research. Firstly, I only deal with the system of single input and single output
in this chapter. For monetary policy evaluation, I have to impose restrictions so
that the usual two-equation system consisted of the IS and Phillips curves can
be analyzed within the current framework. It is appealing to extend the existing
results to bivariate or multivariate cases. To work out this extension, I need to cope
with the transfer matrix and spectral density matrix instead of the scalar sensitivity
function and spectral density function. Design limits will still show up but in the
matrix form [Brock et al (2008b)]. Thus, some new techniques may be required.

Secondly, my analysis can be extended to a hybrid model containing both forward-looking and backwards-looking elements with some complications. As noted in Section 3.2, expectation is potentially another channel through which measurement error influence policy decision. The nature of design limits also depends on the way in which forward-looking elements determine current macroeconomic state variables; specifically, relative to a purely passive policy baseline, a feedback from expectations of future state variables to current state variables is necessary for the existence of a stabilization rule that reduces variance at all frequencies. This suggests, in considering the interplay between model uncertainty and measurement error, it is meaningful to include forward-looking elements, maybe rational expectation based, into the model.

Finally and maybe most importantly, this chapter is silent to a fundamental question: where does measurement error come from? I simply assume that measurement error is there, attached to the true measure, as most previous literature on the subject does. However, an observation is the aggregation of at least three elements: true value, measurement error, and the error due to model uncertainty. They are not mutually independent. For example, measurement error and model uncertainty are generically connected because we measure variables and form expectations based on models. Model uncertainty has already taken effect in the set of observations before we base our robustness analysis on the observations. This chapter makes progress towards but does not complete the full theory of interaction between measurement error and model uncertainty. There are also
other endogenous sources of measurement error such as demographic change and technological development. A structural model on the origins of measurement error will be necessary for a more complete understanding of the policy implications of its presence.
A  APPENDIX OF CHAPTER 2: DERIVATIONS AND PROOFS
A.1 Existence of Nash Equilibrium in the Single-Child Model

The parent needs to maintain nonnegative investment and consumption. Hence, \( y_c \in [0, w_p] \). Notice that \( e_c \in [0, \hat{e}_c] \). Denote \( \Pi \equiv [0, w_p] \times [0, \hat{e}_c] \), which is a convex compact subset of the Euclidean space \( \mathbb{R}^2 \). Given the Inada conditions, there is no need to consider boundary conditions for the maximization problems of the parent and child. As long as conditions (2.6) and (2.7) are satisfied, the corresponding profile \((y_c^N, e_c^N)\) solves their problems, given the other’s choice, and hence is a Nash equilibrium. Since all functions involved in (2.6) and (2.7) are twice continuously differentiable in open sets \((0, w_p)\) and \((0, \hat{e}_c)\), the Implicit Function Theorem guarantees that implicit functions \( y_c = \phi_1(e_c) \) and \( e_c = \phi_2(y_c) \) given by (2.6) and (2.7) are continuously differentiable. The interior-solution assumption implies that they are continuous at the boundaries as well. That is, function \( \phi : \Pi \rightarrow \Pi \) is continuous in \( \Pi \) and hence has a fixed point, by the Brouwer Fixed Point Theorem. A Nash equilibrium is a fixed point of \( \phi \) in \( \Pi \). Therefore, there exists a Nash equilibrium.

A.2 Proof of Proposition 2.1

Proposition 2.1 follows immediately from (2.8) and (2.9).

A.3 Proof of Proposition 2.2

(i) Consider the family planner’s problem first. Maximizing the instantaneous
family objective function (2.11) obtains the first order conditions

\[
\frac{\partial W}{\partial y_c}(y_c^*, e_c^*) = -U'(e_c^*) + (1 + a) r V' f_y(y_c^*, e_c^*) = 0,
\]

\[
\frac{\partial W}{\partial e_c}(y_c^*, e_c^*) = (1 + a) \left[ r V' f_e(y_c^*, e_c^*) - K'(e_c^*) \right] = 0,
\]

Since \( W(y_c, e_c) \) is strictly concave in \( (y_c, e_c) \), the second order conditions automatically hold and the instantaneous family optimum \( (y_c^*, e_c^*) \) is unique. Further, the strict concavity of \( W(y_c, e_c) \) requires the following matrix of the second order derivatives to be negative definite,

\[
\begin{bmatrix}
\frac{\partial^2 W}{\partial y_c^2} & \frac{\partial^2 W}{\partial y_c \partial e_c} \\
\frac{\partial^2 W}{\partial e_c \partial y_c} & \frac{\partial^2 W}{\partial e_c^2}
\end{bmatrix}
= \begin{bmatrix}
U'' + (1 + a) r (V' f_{yy} + r V'' f_y^2) & (1 + a) r (V' f_{ye} + r V'' f_y f_e) \\
(1 + a) r (V' f_{ey} + r V'' f_e f_y) & (1 + a) \left[ r (V' f_{ee} + r V'' f_e^2) - K'' \right]
\end{bmatrix},
\]

for all \((y_c, e_c) \in \Pi\). A matrix is negative definite if all kth order leading principal minors are negative if k is odd and positive if k is even. Therefore, the determinant of this matrix is positive, which yields

\[
\chi_1'(e_c) = \frac{(1 + a) r (V' f_{ye} + r V'' f_y f_e)}{-U'' - (1 + a) r (V' f_{yy} + r V'' f_y^2)} < \frac{K'' - r (V' f_{ee} + r V'' f_e^2)}{r (V' f_{ey} + r V'' f_e f_y)} = \frac{1}{\chi_2'(y_c)},
\]

given that condition (2.10) holds.

For any \((y_c, e_c) \in \Pi\), it is also easy to verify that in expression (2.8), \( \phi_1'(e_c) \leq \phi_2(y_c) \).
\(\chi'_1(e_c)\) and \(\phi'_2(y_c) = \chi'_2(y_c)\) in (2.9). Then,

\[
\phi'_1(e_c) < \frac{1}{\phi'_2(y_c)}.
\]

Therefore, the reaction functions \(y_c = \phi_1(e_c)\) and \(e_c = \phi_2(y_c)\) only cross once in the decentralized setting. There exists a unique fixed point, which is the Nash equilibrium.

To show that \((y_c^N, e_c^N)\) is Pareto efficient, notice that a Pareto allocation that guarantees at least the same utility for the child as the Nash equilibrium \((y_c^N, e_c^N)\) maximizes the parent’s utility (2.1), subject to constraints (2.3)-(2.5) and the child welfare constraint (2.14). One first order condition (with respect to effort choice \(e_c\)) of this Pareto problem is such that \(e_c = \phi_2(y_c)\), exactly the same as given by condition (2.7) in the Nash solution. If the constraint (2.14) is not binding, then the other first order condition (with respect to investment choice \(y_c\)) is \(y_c = \phi_1(e_c)\), also the same as given by (2.6) in the Nash solution. In this case, the Pareto allocation is just the Nash equilibrium. Otherwise, if the constraint (2.14) is binding, \((y_c^N, \phi_2(y_c^N))\) solves (2.14) with equality and the whole Pareto problem. Therefore, the Nash equilibrium \((y_c^N, e_c^N)\) is Pareto efficient.

(ii) The first part of this proof above has shown that the instantaneous family optimum \((y_c^*, e_c^*)\) is unique. In the Nash equilibrium, the first order conditions (2.6) and (2.7) imply

\[
\frac{f_y(y_c^N, e_c^N)}{f_e(y_c^N, e_c^N)} = \frac{U'(c_p^N)}{aK'(e_c^N)} \geq \frac{U'(c_p^N)}{(1 + a)K'(e_c^N)}.
\]

The marginal rate of transformation does not equal the family’s marginal rate of sub-
stitution as required by the instantaneous family optimality. The Nash equilibrium is not family optimal from this instantaneous view.

Then, the proof shows that \((y_c^N, e_c^N) < (y_c^*, e_c^*)\) by contradiction. Suppose it were not true. There are three cases:

C1se 1. \(y_c^N < y_c^*\), but \(e_c^N \geq e_c^*\). Notice that \((y_c^*, e_c^*)\) satisfies condition (2.13). Then,

\[
K'(e_c^*) = rV'(rf(y_c^*, e_c^*)) f_e(y_c^*, e_c^*) \\
\geq rV'(rf(y_c^N, e_c^*)) f_e(y_c^N, e_c^*) \\
\geq rV'(rf(y_c^N, e_c^N)) f_e(y_c^N, e_c^N) \\
= K'(e_c^N),
\]

where the first inequality holds because of the complementarity condition (2.10) and the second because of the concavity of functions \(V\) and \(f\) in \(e_c\).

Since \(K'' > 0\), \(e_c^* > e_c^N\). This is a contradiction.

C2se 2. \(e_c^N < e_c^*\), but \(y_c^N \geq y_c^*\). Since \((y_c^*, e_c^*)\) satisfies condition (2.12), a similar argument shows that

\[
U'(w_p - y_c^*) > U'(w_p - y_c^N).
\]

The concavity of function \(U\) leads to \(y_c^* > y_c^N\), a contradiction again.

C3se 3. \(e_c^N \geq e_c^*\) and \(y_c^N \geq y_c^*\). The instantaneous optimum \((y_c^*, e_c^*)\) satisfies the first order conditions above. Since \(W\) is concave in \((y_c, e_c)\), it must be true that

\[
\frac{\partial W}{\partial y_c}(y_c^N, e_c^N) = -U'(c_p^N) + (1 + a) rV' f_y(y_c^N, e_c^N) \leq 0,
\]
\[ \frac{\partial W}{\partial e_c}(y_c^N, e_c^N) = (1 + a) \left[ rV'f_c(y_c^N, e_c^N) - K'(e_c^N) \right] \leq 0. \]

However, the Nash equilibrium \((y_c^N, e_c^N)\) also satisfy the reaction equations (2.6) and (2.7). Hence,

\[ \frac{\partial W}{\partial y_c}(y_c^N, e_c^N) > -U'(c_p^N) + arV'f_y(y_c^N, e_c^N) = 0, \]

which is a contradiction. Therefore, \((y_c^*, e_c^*) > (y_c^N, e_c^N)\) and \(f(y_c^*, e_c^*) > f(y_c^N, e_c^N)\).

(iii) Notice that the first order conditions for the dictatorial parent’s decisions are consistent with equations (2.6) and (2.7). Then, both the uniqueness and equivalence follow.

### A.4 Existence of Nash Equilibrium in the Multi-Child Model

A strategy profile lives in the space of \(\Pi^n \equiv ([0, w_p] \times [0, \hat{e}_c])^n\), which is a convex compact subset of the Euclidean space \(\mathbb{R}^{2n}\). Due to the Inada conditions, it is sufficient to consider only interior solutions for both the children and parent. The assumptions imposed on utility and production functions ensure that conditions (2.28) and (2.31) are sufficient for the maximization of the objective functions of the parent and children, respectively. The corresponding profile \((y_c^{i,N}, e_c^{i,N})_{i=1}^n\) is a Nash equilibrium. Notice that the implicit function determined by equation (2.28)
can be denoted as \( y_i^c = \phi_i^1(e_i^c, y_i^c) \), while child \( i \)'s reaction function implied by equation (2.31) as \( e_i^c = \phi_i^2(y_i^c, e_i^{-i}) \). Clearly, a Nash equilibrium is a fixed point of function \( \phi : \Pi^n \to \Pi^n \), where \( \phi \equiv [\phi_1^1, ..., \phi_1^n, \phi_2^1, ..., \phi_2^n]' \). As in Appendix A.1, function \( \phi \) is continuous in set \( \Pi^n \) and the Brouwer Fixed Point Theorem establishes the existence of a fixed point.

### A.5 Proof of Proposition 2.3

(i) In the symmetric environment, for any given investment \( y_c \in [0, w_p] \), the cooperative choice of effort \( \phi_2^C \) satisfies equation (2.33), while the noncooperative \( \phi_2^N \) satisfies (2.32). It is convenient to verify \( \phi_2^C(y_c) > \phi_2^N(y_c) \) by contradiction. Suppose the opposite \( \phi_2^C(y_c) \leq \phi_2^N(y_c) \). Recall the assumption that \( S_2 > 0 \) and the concavity of function \( S \). In particular, condition \( S_{11}S_{22} - S_{12}^2 > 0 \) gives that \( S_{11} + S_{22} + 2S_{12} < 0 \). Then, for any \( y_c \), equation (2.33) implies that

\[
K' (\phi_2^C(y_c)) = rV'(w_c^C) f_c (y_c, \phi_2^C(y_c)) + S_1 (\phi_2^C(y_c), \phi_2^C(y_c)) \\
+ S_2 (\phi_2^C(y_c), \phi_2^C(y_c)) \\
\geq rV'(w_c^C) f_c (y_c, \phi_2^C(y_c)) + S_1 (\phi_2^N(y_c), \phi_2^N(y_c)) \\
+ S_2 (\phi_2^N(y_c), \phi_2^N(y_c)) \\
\geq rV'(w_c^N) f_c (y_c, \phi_2^N(y_c)) + S_1 (\phi_2^N(y_c), \phi_2^N(y_c)) \\
+ S_2 (\phi_2^N(y_c), \phi_2^N(y_c)) \\
\geq rV'(w_c^N) f_c (y_c, \phi_2^N(y_c)) + S_1 (\phi_2^N(y_c), \phi_2^N(y_c)) \\
= K' (\phi_2^N(y_c))
\]
The convexity of disutility function $K$ implies that $\varphi_2^C(y_c) > \varphi_2^N(y_c)$, a contradiction.

(ii) Differentiating equation (2.33),

$$\varphi_2^C'(y_c) = \frac{r(V'f_{ye} + rV''f_y f_e)}{K'' - r(V'f_{ee} + rV''f_y^2) - (S_{11} + S_{22} + 2S_{12})}.$$  

Similarly, differentiating equation (2.34) gives

$$\varphi_1'(e_c) = \frac{ar(V'f_{ye} + rV''f_y f_e)}{ar(V'f_{yy} + rV''f_y^2) + nU''}.$$  

When condition (2.10) holds, both $\varphi_2^C'(y_c) > 0$ and $\varphi_1'(e_c) > 0$. Condition $\varphi_1'(e_c)\varphi_2^C'(y_c) < 1$ then requires that the inverse function of $\varphi_2^C(y_c)$ is steeper than function $\varphi_1(e_c)$ at any $e_c \in [0, \hat{e}_c]$. Since $\varphi_2^C(y_c) > \varphi_2^N(y_c)$ for any $y_c$, the following must be true

$$0 < \varphi_1'(e_c) < \frac{y_c^C - y_c^N}{e_c^C - e_c^N} < \frac{1}{\varphi_2^C'([\varphi_2^C]^{-1}(e_c^C))}.$$  

Hence, either $(y_c^C, e_c^C) > (y_c^N, e_c^N)$ or $(y_c^C, e_c^C) < (y_c^N, e_c^N)$. However, since $\varphi_2^C(y_c) > \varphi_2^N(y_c)$ and $\varphi_1'(e_c) > 0$, $(y_c^C, e_c^C) < (y_c^N, e_c^N)$ is impossible. Accordingly, it must be $(y_c^C, e_c^C) > (y_c^N, e_c^N)$ in equilibrium.

On the other hand, when condition (2.15) holds, $\varphi_2^C'(y_c) < 0$ and $\varphi_1'(e_c) < 0$. This must give the following:

$$0 > \varphi_1'(e_c) > \frac{y_c^C - y_c^N}{e_c^C - e_c^N} > \frac{1}{\varphi_2^C'([\varphi_2^C]^{-1}(e_c^C))}.$$
Then conditions that $\varphi_C^C(y_c) > \varphi_N^N(y_c)$ for any $y_c \in [0, w_p]$ and that $\varphi'_1(e_c) < 0$ guarantee $e^C_c > e^N_c$ and $y^C_c < y^N_c$. Whether the human capital output $H_c$ is higher or lower with the cooperative decision than with the noncooperative is left undetermined.

### A.6 Proof of Proposition 2.4

In the symmetric setting, functions $f^i(\cdot, \cdot)$ are the same across $i$, denoted as $f(\cdot, \cdot)$. Differentiating the first order condition (2.31) with respect to $\bar{e}^{-i}_c$ obtains

$$
\frac{de^i_c}{d\bar{e}^{-i}_c} = \frac{r (V'f_{ey} + r V''f_e f_y) \frac{dy^i_c}{de^i_c} + S_{12}}{K'' - S_{11} - r (V'f_{ee} + r V''f^2_e)}.
$$

When condition (2.35) holds, it must be $\frac{de^i_c}{d\bar{e}^{-i}_c} < 0$, or formally, $\frac{d\varphi^2_2(y^i_c, \bar{e}^{-i}_c)}{d\bar{e}^{-i}_c} < 0$. Then the reaction curve $e^i_c = \varphi^2_2(y^i_c, \bar{e}^{-i}_c)$ will cross the 45 degree line and only once. There is a unique fixed point $e_c = \varphi^2_2(y_c, e_c)$ and hence a unique symmetric equilibrium.

### A.7 Proof of Proposition 2.5

(i) As assumed, only consider interior solutions. Given the linear specification of this multi-child model, the optimal condition for the parent’s decision is

$$
y^i_c = \frac{\alpha_1 r A^i_c - \theta_1}{\theta_2 + \alpha_2} + \frac{\theta_2 (w_p - \sum_{j \neq i} y^j_c)}{\theta_2 + \alpha_2}.
$$

Aggregating across $i$ obtains the parent’s average human capital investment $\bar{y}_c$ as
given by (2.38).

Child $i$'s first order condition of optimal decision making is

$$\eta_1 r \kappa y_c^i - J(e_c^i - \bar{e}_c^{-i}) - \alpha - e_c^i = 0,$$

with group utility specification (2.36), and

$$\eta_1 r \kappa y_c^i + J\bar{e}_c^{-i} - \alpha - e_c^i = 0,$$

with specification (2.37). Aggregating them across $i$ gives equations (2.39) and (2.40), respectively.

(ii) Notice that $\bar{y}_c$ is a linear function of $\bar{e}_c$ (contained in $\bar{A}_c$) in (2.38), and so is $\bar{e}_c$ of $\bar{y}_c$ in (2.39) or (2.40). Two choice functions can intersect once and only once, which gives a unique fixed point in $\mathbb{R}^2$. Since Proposition 2.5 assumes that an interior solution $(\bar{y}_c^N, \bar{e}_c^N)$ exists, then it must be unique. Under specification (2.36), equilibrium $(\bar{y}_c^N, \bar{e}_c^N)$ is jointly determined by (2.38) and (2.39). But the measure of social influence $J$ does not enter into these conditions. Hence, $(\bar{y}_c^N, \bar{e}_c^N)$ is independent of $J$.

Under (2.37), the equilibrium is determined by (2.38) and (2.40). To solve for $(\bar{y}_c^N, \bar{e}_c^N)$, insert (2.38) into (2.40):

$$[(1 - J)e_c + \alpha] (n\theta_2 + a\eta_2) = \eta_1 r \kappa L + a(\eta_1 r \kappa)^2 \bar{e}_c.$$
where \( L \equiv -\theta_1 + \theta_2 w_p + a\eta_1 r (\rho \bar{a} + \sigma H_p + \sum_h \beta_h \bar{x}_h) \). Then, immediately,

\[
\bar{e}_c^N = \frac{\eta_1 r \kappa L - \alpha (n\theta_2 + a\eta_2)}{(1 - J)(n\theta_2 + a\eta_2) - a(\eta_1 r \kappa)^2}.
\]

Take derivative with respect to \( J \):

\[
\frac{\partial \bar{e}_c^N}{\partial J} = \frac{[\eta_1 r \kappa L - \alpha (n\theta_2 + a\eta_2)](n\theta_2 + a\eta_2)}{((1 - J)(n\theta_2 + a\eta_2) - a(\eta_1 r \kappa)^2)^2}.
\]

According to (2.38),

\[
\frac{\partial y_c^N}{\partial J} = \frac{a\eta_1 r \kappa}{n\theta_2 + a\eta_2} \frac{\partial \bar{e}_c^N}{\partial J},
\]

which has the same sign as \( \frac{\partial \bar{e}_c^N}{\partial J} \). The interior solution is such that \( \bar{e}_c^N > 0 \). Therefore, when inequality \( \sigma > 1 - J \) holds, \( \bar{e}_c^N > 0 \) will need \( \eta_1 r \kappa L - \alpha (n\theta_2 + a\eta_2) > 0 \), which implies \( \frac{\partial \bar{e}_c^N}{\partial J} < 0 \) and hence \( \frac{\partial y_c^N}{\partial J} < 0 \). In contrast, if instead inequality \( 1 - J > \sigma \) holds, the same argument obtains \( \frac{\partial \bar{e}_c^N}{\partial J} > 0 \) and \( \frac{\partial y_c^N}{\partial J} > 0 \).

### A.8 Existence of Nash Equilibrium and Potential

#### Uniqueness of Symmetric Equilibrium in the Neighborhood Interactions Model

To show the existence of a Nash equilibrium \((y_{c}^{i,N}, e_{c}^{i,N})_{i \in g}\) of this model, notice that both \( y_{c}^{i} = \phi_{1}^i (e_{c}^{i}) \) and \( e_{c}^{i} = \phi_{2}^i (y_{c}^{i}, \bar{e}_c^g) \) are continuous for any \( i \in g \), given the
assumptions on preferences and technologies. Rewrite

\[ e^i_c = \phi_2^i(\phi_1^i(e^i_c), \bar{g}_c^i), \]

which solves \(e^i_c\) as a function of \(\bar{g}_c^i\), \(e^i_c = \xi^i(\bar{g}_c^i)\). Take average:

\[ \bar{g}_c^i = \int_{i \in g} \xi_i^i(\bar{g}_c^i) \, di. \]

A fixed point \(\bar{g}_c^i\) exists in set \([0, \hat{e}_c]\) since functions \(\xi_i^i\) are continuous (by the Implicit Function Theorem). Once \(\bar{g}_c^i\) is solved, all \(e^i_N, i \in g\), are determined by \(e^i_c = \xi^i(\bar{g}_c^i)\), and all \(y^i_N\) by \(y^i_c = \phi_1^i(e^i_c)\). This proves the existence of a Nash equilibrium \((y^i_N, e^i_N)_{i \in g}\).

With identical families in the community, there is no loss of generality to consider only symmetric equilibria \((y^{i,S}_c, e^{i,S}_c)_{i \in g}\). It is defined by relations \(y^{i,S}_c = \phi_1^i(e^{i,S}_c)\) and \(e^{i,S}_c = \phi_2^i(y^{i,S}_c, e^{i,S}_c)\). Rewrite them together into one equation

\[ e^i_c = \phi_2^i(\phi_1^i(e^i_c), \bar{g}_c^i). \]

Then,

\[ \frac{de^i_c}{d\bar{g}_c^i} = \frac{\phi_{2e}}{1 - \phi_{2y} \phi_{1e}}. \]

Since \(e^i_c = \phi_2^i(y^i_c, \bar{g}_c^i)\) is defined by (2.44), the strategic complementarity assumption implies that

\[ \phi_{2e} = \frac{S_{12}}{K'' - S_{11} - \tau(V'f_{ee} + rV''f_e^2)} > 0. \]
When inequality $\phi_2 y \phi_1 > 1$ holds, $\frac{de_i}{de_{\tilde{e}_c}} < 0$ and hence the fixed point $e^S_c = \phi_2(\phi_1(e^S_c), e^S_c)$ is unique. Therefore, there is a unique symmetric equilibrium in this case.

### A.9 Proof of Proposition 2.6

Plug the decision rule of parents (2.45) into children’s reaction functions (2.46) and (2.47), respectively, and solve for $e^i_c$. Equations (2.48) and (2.49) follow immediately. Averaging (2.48) and (2.49) across children obtains (2.50) and (2.51), respectively. It is evident from (2.50) that $\bar{e}_g$ is independent of $J$ and from (2.51) that it is increasing in $J$ if $-\alpha + \bar{B} > 0$. Notice that $\bar{e}_g > 0$ implies that $-\alpha + \bar{B} > 0$ only when $1 - J > 0$.

### A.10 Proof of Proposition 2.7

(i) When social component $S(e^i_c, \bar{e}_g)$ is absent from children’s utility, their decisions in period 2 only respond to parental investment, $e^i_c = \phi_1(y^i_c)$, i.e.,

$$rf_c^i(y^i_c, e^i_c) = \frac{K'(e^i_c)}{V'(w^i_c)}$$

for any $y^i_c \in [0, \bar{w}_p^i]$.

In period 2, parental investments are determined by $y^i_c = \phi_1(e^i_c, \bar{w}_p^{i,g})$ across families, or implicitly,

$$U'(\bar{w}_p^{i,g} - y^i_c) = arV'(w^i_c)f^i(y^i_c, e^i_c),$$

for any $e^i_c \in [0, \hat{e}_c]$. 
Decisions of parents and children, \(y_c^{i,N}(\tilde{w}_p^{i,g})\) and \(e_c^{i,N}(\tilde{w}_p^{i,g})\), then both depend on net family wealth \(\tilde{w}_p^{i,g} \equiv w_p^{i} - \rho_g\). Denote the optimized utility of the parents under net wealth \(\tilde{w}_p^{i,g}\) as

\[
W(\tilde{w}_p^{i,g}) \equiv U(\tilde{w}_p^{i,g} - y_c^{i,N}(\tilde{w}_p^{i,g})) \\
+ a \left[ V(y_c^{i,N}(\tilde{w}_p^{i,g}), e_c^{i,N}(\tilde{w}_p^{i,g})) - K(e_c^{i,N}(\tilde{w}_p^{i,g})) \right].
\]

The problem of residential decision in period 1 is:

\[
\max_{g \in \{A, B\}} W(\tilde{w}_p^{i,g}).
\]

By the Envelope Theorem,

\[
\frac{dW(\tilde{w}_p^{i,g})}{d\tilde{w}_p^{i,g}} = U'(\tilde{w}_p^{i,g} - y_c^{i,N}(\tilde{w}_p^{i,g})) > 0.
\]

Therefore, the residential decision for \(i\) is

\[
g = \begin{cases} 
A & \text{if } \rho_A < \rho_B, \\
B & \text{if } \rho_B < \rho_A, \\
A \text{ or } B & \text{if } \rho_A = \rho_B.
\end{cases}
\]

Since the above argument is true for both \(i = h\) and \(l\), an equilibrium must have \(\rho_A = \rho_B\) to clear the housing market. Further, when housing prices are equal between the two communities, it is evident from the expression of \(W(\tilde{w}_p^{i,g})\) that family \(i\) is indifferent them, either \(i = h\) or \(l\). Therefore, any community configuration can
emerge in an equilibrium.

(ii) When child effort is the only factor of human capital production,

\[ H^i_c = f^i_c(e^i_c), \]

effort inputs of all children are jointly determined by \( e^{i,g}_c = \phi^i_c(0, \bar{e}^g_c), \) i.e.,

\[ rf^i_c(e^i_c) = \frac{K'(e^i_c) - S_1(e^i_c, \bar{e}^g_c)}{V'(w^i_c)}, \] for all \( i \in g. \)

And there is no parental investment. Therefore, the residential decision problem in period 1 is

\[ \max_{g \in \{A, B\}} U(w^i_p - \rho_g) + a \left[ V(rf^i(e^{i,g}_c)) - K(e^{i,g}_c) \right]. \]

Since all children are ex ante identical, it must be true that \( \bar{e}^A_c = \bar{e}^B_c, \) which also implies that \( e^{i,A}_c = e^{i,B}_c. \) Then the residential choice for \( i \) is just the same as in case (i), no matter \( i = h \) or \( l. \) Condition \( \rho_A = \rho_B \) holds to clear the market, and any community configuration is compatible with an equilibrium.

### A.11 Proof of Proposition 2.8

In the current setting, function \( W(t, \rho, x) \) takes the form of

\[
W(t, \rho, x) = U(\tilde{w}^{i,g}_p - y^{i,N}_c(\tilde{w}^{i,g}_p, x)) + aV\left(rf\left(y^{i,N}_c(\tilde{w}^{i,g}_p, x), e^{i,N}_c(\tilde{w}^{i,g}_p, x)\right)\right) + aS\left(e^{i,N}_c(\tilde{w}^{i,g}_p, x), \bar{e}^{i,N}_c(\tilde{w}^{i,g}_p, x)\right) - aK\left(e^{i,N}_c(\tilde{w}^{i,g}_p, x)\right).
\]
By the Envelope Theorem,

\[ W_x(t, \rho, x) = aS_2 \frac{\partial \bar{e}_c^g}{\partial x}, \]

\[ W_\rho(t, \rho, x) = -\frac{\partial W(t, \rho, x)}{\partial \tilde{w}_i^ \rho} + aS_2 \frac{\partial \bar{e}_c^g}{\partial \rho} = -U' + aS_2 \frac{\partial \bar{e}_c^g}{\partial \rho}. \]

Substituting them into (2.52), equation (2.54) follows.

Take derivative of \( R(t, \rho, x) \) in expression (2.54) with respect to \( t \), and notice that either type of family faces the same \( \bar{e}_g^c \) in neighborhood \( g \), which implies that \( \frac{\partial \bar{e}_c^g}{\partial t} = 0, \frac{\partial^2 \bar{e}_c^g}{\partial \rho \partial t} = 0, \) and \( \frac{\partial^2 \bar{e}_c^g}{\partial x \partial t} = 0 \). Therefore, \( R_t(t, \rho, x) \) reduces to

\[ R_t(t, \rho, x) = \frac{a \frac{\partial \bar{e}_c^g}{\partial x} \left( S_{12} U' \frac{\partial e_i^g}{\partial t} - S_2 \frac{\partial U'}{\partial t} \right)}{\left( U' - aS_2 \frac{\partial \bar{e}_c^g}{\partial \rho} \right)^2} . \]

Notice that \( \frac{\partial \bar{e}_c^g}{\partial x} \) has the same sign as \( \frac{\partial e_i^g}{\partial t} \). When \( \frac{\partial e_i^g}{\partial t} > 0 \), condition \( S_{12} > S_2 \frac{\partial U'/\partial t}{\bar{e}_c^g/\partial t} \) implies that \( S_{12} U' \frac{\partial e_i^g}{\partial t} - S_2 \frac{\partial U'}{\partial t} > 0 \) and \( R_t(t, \rho, x) > 0 \). When \( \frac{\partial e_i^g}{\partial t} < 0 \), the condition implies that \( S_{12} U' \frac{\partial e_i^g}{\partial t} - S_2 \frac{\partial U'}{\partial t} < 0 \) but still \( R_t(t, \rho, x) > 0 \). By Lemma 2.1, the rest of Proposition 2.8 follows.

**A.12 Proof of Proposition 2.9**

Recall that \( \bar{e}_c^g \) is still given by (2.50) and (2.51) under social utility functions (2.36) and (2.37), respectively, except that individual abilities \( a_i \) and other characteristics \( x_h \) are ex ante identical. Also notice that wealth \( \tilde{w}_i^ {i,g} \) are not in the expression of \( B_i \).
Then, in the two cases,

\[ \frac{\partial \bar{e}_g}{\partial x} = \frac{K}{1-\sigma} \quad \text{and} \quad \frac{\partial \bar{e}_c}{\partial x} = \frac{K}{1-J-\sigma}. \]

Similarly, since housing price \( \rho_g \) is subtracted from total wealth endowment \( w_p \),

\[ \frac{\partial \bar{e}_c}{\partial \rho} = -\frac{H}{1-\sigma} \quad \text{and} \quad \frac{\partial \bar{e}_g}{\partial \rho} = -\frac{H}{1-J-\sigma}. \]

Function (2.52) takes the form of (2.56) and (2.57) in these cases, respectively.

Furthermore, in the linear quadratic model,

\[ U' = \theta_1 - \theta_2 (w_p^i - \rho_g - y_i^{i,N}). \]

From equation (2.45) (with net wealth in) and the fact that effort input (2.48) or (2.49) is also linear in net wealth, it is straightforward that, in both cases,

\[ 0 < U'|_{t=h} < U'|_{t=1}. \]

In the case of social utility function (2.36),

\[ S_2 = J(e_c^i - \bar{e}_c^g). \]

Due to (2.48), it must be true that \( e_c^h > e_c^l \). Then,

\[ S_2|_{t=h} > 0 > S_2|_{t=1}. \]
Therefore, $R(t, \rho, x)$ is increasing in $t$ in this case, $R(h, \rho, x) > R(l, \rho, x)$. And, in the case of social utility function (2.37),

$$S_2 = Je^i_c.$$ 

Since $e^h_c > e^l_c$ still holds, it is then obvious that

$$S_2|_{t=h} > S_2|_{t=l}.$$ 

Hence, $R(t, \rho, x)$ is again increasing in $t$. According to Lemma 2.1, therefore, Proposition 2.9 holds.
B  APPENDIX OF CHAPTER 3: SOME EXAMPLES AND RESULTS
B.1 Proof of Bode’s Constraint (3.10)

This proof of Bode’s constraint (3.10) is an application of Wu and Jonchheere’s (1992) lemma, which states that

$$\int_{-\pi}^{\pi} \ln |e^{i\omega} - r|^2 d\omega = \begin{cases} 0 & \text{if } |r| \leq 1, \\ 2\pi \ln |r|^2 & \text{if } |r| > 1. \end{cases}$$

Factorize both the numerator and denominator of the sensitivity function defined in (3.7) by the fundamental theorem of algebra,

$$|S(\omega)| = \left| \frac{1 - A(\omega)e^{-i\omega}}{1 - A(\omega)e^{-i\omega} + B(\omega)F(\omega)e^{-i\omega}} \right| = \prod_{j=1}^{m} \left| 1 - \lambda_j e^{-i\omega} \right| \prod_{j=1}^{n} \left| 1 - \mu_j e^{-i\omega} \right|,$$

where $$\lambda_j$$ and $$\mu_j$$ are open-loop and closed-loop poles of the system. Then,

$$\int_{-\pi}^{\pi} \ln(|S(\omega)|^2) d\omega = \sum_{j=1}^{m} \int_{-\pi}^{\pi} \ln |e^{i\omega} - \lambda_j|^2 d\omega - \sum_{j=1}^{n} \int_{-\pi}^{\pi} \ln |e^{i\omega} - \mu_j|^2 d\omega$$

$$= 4\pi \sum_{j:|\lambda_j|>1} \int_{-\pi}^{\pi} \ln |\lambda_j| d\omega$$

$$\geq 0.$$

The second equality uses Wu and Jonchheere’s lemma, and it follows because all closed-loop poles $$\mu_j$$ are inside the unit disk in the complex plane; otherwise, the controlled system is not stationary and hence has not been stabilized by the control. The last inequality binds when all $$\lambda_j$$ are inside the unit disk, i.e. when the uncontrolled system is stationary, and is strict when some $$\lambda_j$$ are outside the unit.
Bode’s constraint (3.10) holds.

B.2 Real Sensitivity Function of Optimal Policy Rule with Nonfiltered Data

This section provides some further intuitions for the real sensitivity function of the optimal policy rule with nonfiltered data developed in Section 2.3. As evident from expression (3.9), a good control should make the modulus of both $S(\omega)$ and $T(\omega)$ as small as possible at each $\omega$.

Consider two sensitivity functions, $\tilde{S}(\omega)$ and $S(\omega)$, and focus on frequency $\omega$, where $\tilde{S}(\omega)$ is complex and $S(\omega)$ is real, but the modulus of $\tilde{S}(\omega)$ is the same as $S(\omega)$. As shown in Figure (B.1), $S(\omega)$ and $\tilde{S}(\omega)$ are then on the same circle in the complex plane. Due to the complementarity principle,

$$\tilde{T}(\omega) + \tilde{S}(\omega) = 1 \text{ and } T(\omega) + S(\omega) = 1.$$

Figure (B.1) shows that the angle $C$ is always obtuse, the modulus of $\tilde{T}(\omega)$ is always greater than that of $T(\omega)$. Therefore, $S(\omega)$ performs better than $\tilde{S}(\omega)$ at frequency $\omega$.

This argument follows at all frequencies, so the optimal solution must set the sensitivity function to be real. As argued in the text, a complex sensitivity function induces indelible phase shifts to the measurement error process $n_t$, which causes
Figure B.1: Complex versus Real Sensitivity Functions, at Frequency $\omega$

Note: At frequency $\omega$, $S(\omega)$ is real, $\tilde{S}(\omega)$ is complex, $|S(\omega)| = |\tilde{S}(\omega)|$, $S(\omega) + T(\omega) = 1$, and $\tilde{S}(\omega) + \tilde{T}(\omega) = 1$.

side noise effects in the presence of data noise.

## B.3 Optimal Policy Rule with Nonfiltered Data in the Absence of Data Noise

The optimal policy rule $S^E(\omega)$ with nonfiltered data is characterized by (3.28) with spectral target $\lambda^E$ determined by (3.29). For general case, I have to numerically parameterize the model and seek a quantitative solution. I use the bisection method for numerical implementation.\(^1\)

\(^1\)Brock and Durlauf (2005) used Jensen’s inequality to recover the representation of the feedback rule in the absence of measurement errors.
In the case absent of data noise, however, I can show that the solution \( S^E(\omega) \) reduces to the benchmark policy \( S^B(\omega) \). Notice that \( f_n(\omega) = 0, \forall \omega \) in this case and hence equation (3.29) gives

\[
\ln \lambda^E = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln (f_{x^{nc}}(\omega)) \, d\omega.
\]

Given (3.5), we have \( f_{x^{nc}}(\omega) = \frac{\sigma^2_\varepsilon}{2\pi} \left| \frac{W(\omega)}{1 - \Lambda(\omega)e^{-i\omega}} \right|^2 \). By factorizing \( f_{x^{nc}}(\omega) \) and applying Wu and Jonchheere’s result,

\[
\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln (f_{x^{nc}}(\omega)) \, d\omega = \ln \left( \frac{\sigma^2}{2\pi} \right).
\]

Therefore, \( \lambda^E = \lambda^N \equiv \frac{\sigma^2}{2\pi} \), which is same as the noise-free target. Substituting \( \lambda^E = \frac{\sigma^2}{2\pi} \) into equation (3.28), it follows that

\[
|S^E(\omega)|^2 = \frac{\sigma^2_\varepsilon}{2\pi f_{x^{nc}}(\omega)},
\]

with is identical to \( |S^B(\omega)|^2 \) in equation (3.14).

### B.4 Policy Rules without Lagged Terms

In the monetary model described in Section 4, the policymaker observes that the loss function (3.56) takes the form of

\[
V(\pi) = \frac{(b\beta)^2 \sigma^2_n + \sigma^2_\varepsilon}{1 - (\rho - b\beta)^2},
\]
under the control of a policy rule without lagged terms as (3.57). To minimize the loss function, he derives the first-order conditions

$$b\beta \left[1 - (\rho - b\beta)^2\right] \sigma_n^2 = [(b\beta)^2\sigma_n^2 + \sigma_\varepsilon^2] (\rho - b\beta).$$

There are two roots to this equation. One is $\beta^E$ as given in (3.59). Note that $|\rho - b\beta| < 1$ to ensure that $\pi_t^E$ is stationary under the control. The other root violates this requirement since $|\rho| < 1$, and is not a solution.

Policy rule $\beta^N$ in (3.58) follows obviously. Then,

$$|\beta^E| \leq |\beta^N| \Leftrightarrow \frac{1}{2|\rho|\sigma_n^2} \left[\sqrt{[(1 - \rho^2)\sigma_n^2 + \sigma_\varepsilon^2]^2 + 4\rho^2\sigma_n^2\sigma_\varepsilon^2 - [(1 - \rho^2)\sigma_n^2 + \sigma_\varepsilon^2]} \leq |\rho| \right] \Leftrightarrow \sqrt{[(1 - \rho^2)\sigma_n^2 + \sigma_\varepsilon^2]^2 + 4\rho^2\sigma_n^2\sigma_\varepsilon^2} \leq \sigma_n^2 + \sigma_\varepsilon^2 + \rho^2\sigma_n^2 \Leftrightarrow 0 \leq 4\rho^2\sigma_n^4,$$

which proves the comparison result in (3.60).

### B.5 Wiener Filter in the AR (1) Model

This section gives more details on the Wiener Filter in the AR (1) model that is used in Section 4.3. To start, notice that for AR (1) process $\pi_t$,\n
$$f_{\pi}(z) = \frac{\sigma_\varepsilon^2}{2\pi(1 - \rho z)(1 - \rho z^{-1})}.$$\n
Equation (3.66) then is based on this result.
The following shows why equation (3.69) holds:

\[
\begin{align*}
\left[ \frac{f_\pi(z)}{H^*_\pi(z)} \right]_+ &= \frac{\sigma_z^2}{\sqrt{2\pi k}} \left[ \frac{1}{(1 - \rho z)(1 - \Phi z^{-1})} \right]_+ \\
&= \frac{\sigma_z^2}{(1 - \rho \Phi)\sqrt{2\pi k}} \left[ \frac{1}{1 - \rho z} + \frac{\Phi}{z - \Phi} \right]_+ \\
&= \frac{\sigma_z^2}{(1 - \rho \Phi)(1 - \rho z)\sqrt{2\pi k}},
\end{align*}
\]

where the last equality follows by observing that the term \(\frac{\Phi}{z - \Phi}\) has a pole at \(z = \Phi\) inside the unit circle and hence is a forward transform which can be ignored.

Given (3.69), the causal Wiener filter is given as

\[
M(z) \equiv \left[ \frac{f_\pi(z)}{H^*_\pi(z)} \right]_+ = \frac{\Phi \sigma_z^2}{(1 - \rho \Phi)(1 - \Phi z)\rho \sigma_n^2},
\]

which has the representation in the frequency domain as (3.70).


