Essays on Social Learning, Cooperation, Asset Markets and Human Capital.

James Best

Doctor of Philosophy

School of Economics
University of Edinburgh
2014

To my family,

It would have been impossible without you.

0.1 Declaration of Own Work

I declare that this thesis was written and composed by myself and is the result of my own work unless clearly stated and referenced. Chapter 2 of this thesis is co-authored with Prof. John Moore of the University of Edinburgh and London School of Economics. I made substantial contributions to the second chapter. This thesis has not been submitted for any other degree or professional qualifications.

James Best

0.2 Acknowledgements

I am deeply indebted to many people. My supervisor Prof. Ed Hopkins' patient support, friendship, and advice have been truly invaluable. My co-author Prof. John Moore has been a great mentor and friend to me. I feel that few have had the good fortune that I have had to be able to work so closely with someone of John's brilliance. Jozsef Sakovics, also my supervisor, has always had an open door and a wise remark whenever I have needed to discuss new ideas and problems. In general, the faculty at the Edinburgh School of Economics have been immensely supportive at all points and I feel very fortunate and privileged to have had the opportunity to pursue my research at this School. Likewise, the people at the Scottish Graduate Programme in Economics have been equally supportive.

I am also indebted to those who I worked with at the University of Chicago . Particularly to Professors James Heckman, Steven Durlauf and Luigi Zingales. Regarding Professor James Heckman, it would be difficult to overstate the value of just having the opportunity to learn from someone of his genius. Professor Heckman did more than this as he also gave me the opportunity to stay in Chicago and conduct research on the implications of genes and environment for human capital policy. This subject had long fascinated me but it had never seemed viable as it was so far from my research agenda at the time. Having Prof. Heckman's support and guidance gave me the courage to tackle such an important and difficulty problem. Prof. Durlauf is a brilliant scientist and an inspiring instructor. He gave me great support in my research agenda and also provided me with much needed support through some tough times. Without Prof. Luigi Zingales I would have never been in Chicago. He invited me to study at the Booth School of Business, provided me with the funding to do so and gave me invaluable experience working as his research assistant.

I owe many of my fellow PhD students for their friendship and advice. Particularly, Drs. Sean Brocklebank, David Comerford, Thomas Flochel and (soon to be Dr.) David Pugh of the University of Edinburgh; Keshav Dogra of Columbia University; and Cullen Roberts of the University of Chicago.

Regarding chapter 1, I am grateful to Ed Hopkins, Jozsef Sakovics, Jakub Steiner, Keshav Dogra, Tatiana Kornienko, Stephen Morris, Stephen Durlauf, Philip Reny, Kohei Kawamura, Jonathan Thomas, David Pugh, Sean Brocklebank, Nick Vikander, and Jose V. Rodriguez Mora for their comments and criticisms on earlier draughts of this paper. Regarding chapter 2, I am most grateful to my co-author John Moore. Also to Ed Hopkins, Tatiana Kornienko, Jozsef Sakovics, Jean Tirole, Andrew Clausen, Joel Sobel, Thomas Marriotti, Keshav Dogra, David Pugh and Sean Brocklebank for their comments and criticisms. Regarding chapter 3, I am most grateful to James Heckman for guiding my research agenda and acting as my mentor. Likewise, to Steven Durlauf who has played a large role in the development of this chapter. Also, I would like to thank Ed Hopkins, Jakub Steiner, Andrew Clausen, Michele Bélot, Philipp Kircher, Dimitra Politi, Steven Dieterle, Mike Elsby, Keshav Dogra, David Pugh and Sean Brocklebank for their comments and criticisms.

This work was produced as a post-graduate student at the University of Edinburgh and as a visiting student at the University of Chicago. I received funding from the Edinburgh School of Economics, the British Economic Social Research Council postgraduate studentship funding scheme and the Scottish Institute for Research in Economics.

0.3 Abstract

In the first chapter, I examine the effect of social learning on social norms of cooperation. To this end I develop an 'anti-social learning' game. This is a dynamic social dilemma in which all agents know how to cooperate but a proportion are informed and know of privately profitable but socially costly, or uncooperative, actions. In equilibrium agents are able to infer, or learn, the payoffs to the actions of prior agents. Agents can then learn through observation that some socially costly action is privately profitable. This implies that an informed agent behaving uncooperatively can induce others to behave uncooperatively when, in the absence of observational learning, they would have otherwise been cooperative. However, this influence also gives informed agents an incentive to cooperate – not cooperating may induce others to not cooperate. I use this model to give conditions under which social learning propagates cooperative behaviour and conditions under which social learning propagates uncooperative behaviour.

In the second chapter, I present a co-authored model of a self-fulfilling price cycle in an asset market. In this model the dividend stream of the economy's asset stock is constant yet price oscillates deterministically even though the underlying environment is stationary. This creates a model in which there is rational excess volatility - 'excess' in the sense that it does not reflect changes in dividend streams and 'rational' in that all agents are acting on their best information. The mechanism that we uncover is driven by endogenous variation in the investment horizons of the different market participants, informed and uninformed.

On even days, the price is high; on odd days it is low.

On even days, informed traders are willing to jettison their good assets, knowing that they can buy them back the next day, when the price is low. The anticipated drop in price more than offsets any potential loss in dividend. Because of these asset sales, the informed build up their cash holdings. Understanding that the market is flooded with good assets, the uninformed traders are willing to pay a high price. But their investment horizon is longer than that of the informed traders: their intention is to hold the assets they purchase, not to resell.

On odd days, the price is low because the uninformed recognise that the informed are using their cash holdings to cherry-pick good assets from the *market. Now the uninformed, like the informed, are investing short-term. Rather than buy-and-hold as they do with assets purchased on even days, on odd days the uninformed are buying to sell.*

Notice that, at the root of the model, there lies a credit constraint. Although the informed are flush with cash on odd days, they are not deep pockets. On each cherry that they pick out of the market, they earn a high return: buying cheap, selling dear. However they don't have enough cash to strip the market of cherries and thereby bid the price up.

The final chapter is on identifying the role of privilege in determining intergenerational mobility. The intergenerational elasticity of income is the standard measurement economists use for intergenerational mobility. It is not clear how we should interpret intergenerational elasticities. Particularly, high intergenerational elasticities could either reflect inequality of opportunity or the importance of genetically heritable characteristics in determining genes. Behavioural geneticists have long been using a twin based variance decomposition method, the ACE model, to estimate the genetic heritability of various characteristics. It is not clear, however, what this approach implies for intergenerational mobility of equality of opportunity.

I develop a novel method that extends the methodology used in behavioural genetics to identifying how much of the intergenerational elasticity of income is determined by the presence (absence) of environmental privileges associated with being children of high (low) earners. Using this approach we can examine the counterfactuals of giving a poorer child the environment of a richer child; equalising the privileges associated with family income; and equalising the family environmental factors not associated with parental income. Furthermore, this method allows us to identify how good parental income is as a measure of family environment. The model I develop nests the behavioural genetics model allowing us to relax some of the identifying assumptions used in the standard ACE model.

Finally, I apply this method to data on the income elasticities between American males of different types of relation: fraternal twins, identical twins and father-son relationships. The results of this application suggest that a 1 percent increase in the privilege associated with parental income increases child income by about 1 tenth of a percent. Equalising, to the mean, the environmental privileges across the population results in about a 30 percent drop in the intergenerational elasticity of income and a 5 percent drop in

the variance of income across the population. These results must be treated tentatively as the twin data comes from a separate survey to the data on intergenerational elasticities.

Contents

Chapter 1

Anti-Social Learning: The Effect of Social Learning on Cooperation.

James Best

I examine the effect of social learning on social norms of cooperation. To this end I develop an 'anti-social learning' game. This is a dynamic social dilemma in which all agents know how to cooperate but a proportion are "informed" and know of privately profitable but socially costly, or uncooperative, actions. In equilibrium agents are able to infer, or learn, the payoffs to the actions of prior agents. Agents can then learn through observation that some socially costly action is privately profitable. This implies that an informed agent behaving uncooperatively can induce others to behave uncooperatively when, in the absence of observational learning, they would have otherwise been cooperative. However, this influence also gives informed agents an incentive to cooperate – not cooperating may induce others to not cooperate. I use this model to give conditions under which social learning propagates cooperative behaviour and conditions under which social learning propagates uncooperative behaviour.

1.1 Introduction

The ability of people to learn from the actions of others frequently propagates socially harmful behaviour. For example, if we see others speeding, shirking at work, or dumping illegally, we may infer that these activities are not effectively punished and engage in them ourselves. Such considerations are not entirely new: Mayor Giuliani used such a 'broken-window' theory (Kelling and Wilson, 1982) of crime to argue for zero-tolerance policing in New York. In these cases people assume that others would not be engaged in these socially harmful activities if they were not privately profitable. Through a similar process people, or organisations, also learn *how* to successfully execute particular socially harmful acts. The criminal methods that criminals learn from one another in prison are an oft lamented example of this. Indeed it has been famously said that "Prisons are universities of crime, maintained by the state" (Kropotkin, 1887). An even starker case is the nuclear weapons programme of the Soviets in the forties and fifties which drew on top secret research acquired, through espionage, from the US and the UK. Similarly, many of the new countries to develop nuclear weapons have done so using the research of earlier nuclear powers.

The purpose of this paper is to clearly illustrate the potential effects of social learning on norms of cooperative, or uncooperative, behaviour. I develop a simple sequential game where some informed agents have independently acquired information about the private profitability of some uncooperative behaviour. In equilibrium, the uncooperative behaviour of informed agents can reveal the profitability of this behaviour to uninformed agents. Consequently, uncooperative acts have a bad influence as they teach uninformed agents the profit in such behaviour, oft resulting in imitation. However, such learning and imitation implies that the original non-cooperators harm themselves, to some extent, by teaching others the profit in some socially harmful activity – they become victims of their own bad influence. If these potential non-cooperators expect to be sufficiently influential they may cooperate, hiding their socially harmful knowledge, in the hope of reducing the harm they suffer from the uncooperative behaviour of others.

A perhaps surprising result emerges: social learning can prevent uncooperative behaviour instead of propagating it. Uncooperative behaviour is prevented when it is unlikely that any one person will independently learn how to profit through uncooperative behaviour. In such circumstances the action of an independently informed agent can be very influential; as other agents will not acquire the information of this agent unless he reveals it through his action. Therefore these informed agents choose to cooperate instead, so that they can prevent the socially harmful behaviour from spreading to others. Here, because of its potential to spread uncooperative behaviour, social learning prevents any uncooperative behaviour happening in the first place. On the other hand, when independent learning is likely, those who are independently informed are not very influential and so do not suppress their uncooperative behaviour. In these cases, the most commonly observed, social learning propagates uncooperative behaviour.

An example may make the argument a little clearer. Blast, or dynamite, fishing is a very effective fishing method. However, blast fishing damages the underlying ecology of the ecosystem and indiscriminately kills large numbers of fish, many of which go to waste (Fox and Erdmann, 2000). Due to the socially costly nature of blast fishing it is now illegal in most countries. However, it was not always illegal, and more importantly, it was not a technique known amongst all fishermen.

Consider now social learning within two different groups of fishermen who have not yet discovered, or begun, the practice of blast fishing. The first group fish on Blue Bay. The second group fish on a large inland lake. These two groups have no direct contact with one another. Blue Bay is one of many bays along a coastline. Fishermen in one of the other bays, Red Bay, have discovered and begun the practice of blast fishing. Most of the Blue Bay fishermen have frequent contact with people from Red Bay. a single lake fisherman has frequent contact with people from Red Bay.

In each group an individual fisherman discovers blast fishing through his contact with people from Red Bay. Both fishermen must then decide whether or not to begin blast fishing. Blast fishing is socially inefficient, so the fishermen would be better off if no one used blast fishing. If the newly informed fishermen begin blast fishing then they will get high fish yields in the short run. However, if they adopt blast fishing then their respective communities will also adopt it, implying lower long run yields. As the other lake fishermen have little contact with people from Red Bay it is unlikely that they will independently learn of blast fishing. If the lake fisherman does not use blast fishing then other lake fishermen will not learn of blast fishing for some time. In contrast, if the bay fisherman does not use blast fishing he only delays the learning process of the other Blue Bay fishermen by a short while. This is because the other Blue Bay fishermen are very likely to learn independently of him; due to their frequent contact with the Red Bay fishermen.

Social learning amongst the lake people causes the lake fisherman to be very influential as his action determines, for quite some time, whether or not people blast fish on the lake. On the other hand, social learning gives the Blue Bay fisherman little informational influence over the people of Blue Bay, because they are so likely to learn independently of him. Given their relative influences it is clear, in the absence of other considerations such as community enforcement, that the lake fisherman should not begin blast fishing while the bay fisherman should. When the probability of independent learning is low then social learning, amongst the lake fishermen, can prevent blast fishing occurring in the first place; and when this probability is high then social learning propagates blast fishing, as with the bay fishermen.

The above example of blast fishing ignores the potential for a group to enforce good behaviour. It may be the case that the bay fisherman does not begin blast fishing because the fishermen are able to collectively enforce cooperative social norms. Theoretical work on repeated games (Mailath and Samuelson, 2006) and community enforcement (Kandori, 1992) demonstrate that cooperation can be sustained by strategies that punish non-cooperators. None the less, in contrast to the bay fisherman, the lake fisherman can ensure a cooperative outcome without needing to worry about such enforcement strategies. This paper examines only finite games so that such punishment strategies are not credible. This allows us to abstract away from issues of social enforcement and isolate the effect of social learning on cooperative behaviour.

Social learning was pioneered in the papers of Banerjee (1992) and Bikhchandani et al. (1992). However, they examine the spread of socially useful information where as this paper examines the spread of socially harmful information and behaviour. Their papers also examines the potential for inefficient aggregation of information. Such issues are not examined in this paper as the aim is to examine the effect of social learning on cooperation and not vice-versa. To this end, those who have information about payoffs in the anti-social learning model below have perfectly accurate information.

Perhaps the closest work to this paper is the work on leadership pioneered by Hermalin (1998). Hermalin shows that group contribution to a public good is higher if information about the value of contribution is restricted to a single known leader acting before all other members of the group. In order to credibly signal the value of contributing to some group project the leader contributes more than he would in the perfect information version of the game. The other members of the group exert the same effort as they would in a perfect information version of the game.

There are several major differences between this paper and Hermalin (1998). First, Hermalin (1998) is a signalling model and the anti-social learning model is not. In this paper informed agents bear the cost of lower private payoffs so they can *conceal* information; whereas, in Hermalin's signalling model the leader bears the cost of higher effort so he can credibly *reveal* information. Second, in Hermalin [1998] there is a single leader who is known to be informed and all other agents are uninformed. The uninformed agents all act after observing the leader's action. In this paper, however, there are many potentially informed agents whose informational status is not known. Last, and most importantly, the anti-social learning model shows how social learning can prevent *or* propagate uncooperative behaviour. Hermalin (1998) does not examine this issue as his paper aims to explain leadership and not the role of social learning in propagating norms of cooperative or uncooperative behaviour.

Acemoglu and Jackson (2011) bears some similarities in its research agenda. They show that the actions of highly observable individuals may help select the equilibrium of a repeated game. In this case learning and expectations help leaders select amongst different equilibria. In the anti-social learning model below, however, people's actions do not serve to select one of many possible equilibria. Instead, if there were no uncertainty or learning in this model then uncooperative behaviour would be the unique dominant strategy. It is shown that introducing uncertainty and social learning into such a game can remove uncooperative behaviour as the dominant strategy equilibrium.

The paper proceeds as follows: the basic model is presented in section II and equilibrium is analyzed in section III. The implications of the equilibrium conditions for the effect of social learning on cooperation is then discussed in section IV. Section V concludes. All proofs not in the main body are presented in the appendix.

1.2 Model

There is a finite population $N = \{1, ..., n\}$. Agents act sequentially in order of their index $i \in N$. They choose an action $a_i \in A = \{c\} \bigcup D$. Action c is a singleton and D is a continuum, $D = [0, 1]$. There is an element $d^* \in D$ drawn from a uniform distribution, $d^* \sim U[0, 1]$. Action c is called *cooperate*; d^* is called *defect*; and any action in $D \backslash \{d^*\}$ is called *blunder*.

Agents have a Von Neumann-Morgenstern utility function $U(a_i, \boldsymbol{a}_{-i}, d^*)$, where $\boldsymbol{a}_{-i}$ denotes the actions of all agents except i and the value of d^* is the state of the world. Each action has a private payoff for i, $v(a_i)$, that is independent of $\boldsymbol{a}_{-i}$. Cooperate, c, has a lower private payoff than defect, d^*, and a higher payoff than blunder. All actions in D have a negative externality attached to them. $X(\boldsymbol{a}_{-i})$ is the total cost to i from the externalities of other agents' actions; referred to in what follows as "the social harm suffered by i". $X(\boldsymbol{a}_{-i})$ is linear in the number of agents playing an action in D and is independent of action a_i.[1] Specifically,

$$U(a_i, \boldsymbol{a}_{-i}, d^*) = v(a_i) - X(\boldsymbol{a}_{-i}) \tag{1.1}$$

where

[1] The results of this paper will be qualitatively similar for more general utility functions where the privately optimal action is socially inefficient. However, a linear additively separable utility function is sufficient for demonstrating the importance of social learning for cooperative behaviour.

$$v(a_i) = \begin{cases} 0 & \text{if } a_i = c \\ \bar{v} > 0 & \text{if } a_i = d^* \\ \underline{v} < 0 & \text{if } a_i \notin \{c, d^*\} \end{cases} \tag{1.2}$$

and,

$$X(\mathbf{a}_s) = \kappa \sum_{i \in s} \mathbb{1}\left[a_i \in D\right] \geq 0 \tag{1.3}$$

for a set of actions $\mathbf{a}_s$.

This game is a social dilemma in the sense that the action with the highest private payoff, d^*, causes greater social harm than the private benefit of that action:

$$\bar{v} < \kappa(n-1). \tag{1.4}$$

That is, all agents defecting is Pareto inferior to all agents cooperating. Note that in the perfect information variant of this game all agents would defect.

The information structure of this game is as follows. Agents independently learn the value of d^* with probability ρ. That is, each agent i receives a signal $s_i \in \{d^*, \emptyset\}$ which is i.i.d. across all agents with a common prior probability of $Pr(s_i = d^*) = \rho$, where $\rho \in (0, 1)$. An agent who independently learns the value of d^* is referred to as *informed*; and an agent who has not is referred to as *uninformed*. Agent i observes the actions of all prior agents, $\boldsymbol{a}_{<i} = \{a_1, ..., a_{i-1}\}$ where $\boldsymbol{a}_{<1} := c$.[2] Agents do not observe whether other agents are informed or uninformed. The information set of i, h_i, is then the history of actions prior to i and i's signal: $h_i = \{\boldsymbol{a}_{<i}, s_i\}$.

1.3 Strategies and Equilibrium

Let the dynamic game defined above be called $G(\rho; n, \{U(.)\}_{i=1}^n)$. I use the term *equilibrium* to mean a (weak) Perfect Bayesian equilibrium of this game (Fudenberg and Tirole, 1991). A pure strategy for player i is a map $\alpha_i : \{A\}^{i-1} \times \{d^*, \emptyset\} \rightarrow A$ where $\alpha_i(\boldsymbol{a}_{<i}, s_i)$ is i's action when i observes history $\boldsymbol{a}_{<i} = \{A\}^{i-1}$ and has received signal $s_i \in \{d^*, \emptyset\}$.

This paper focuses on symmetric pure strategy equilibria. There is, however, an informal discussion of non-symmetric equilibria at the end of this section.

[2] I set $\boldsymbol{a}_{<1} := c$ rather than $\emptyset$ to simplify notation later.

1.3.1 Symmetric Equilibria

Strategies are symmetric if the actions they imply do not substantively alter when the elements of $[0,1]$ are relabelled. This implies that if an informed agent chooses to cooperate (defect) when nature chooses $d^* = x$ they should also choose to cooperate (defect) if $d^* \neq x$.

Define a **permutation function** $f(x)$ to be any function giving a one to one mapping of the union of $\{\emptyset\}$ and A into themselves where $f(\emptyset) = \emptyset$, $f(c) = c$ and $f(x) \in D$ for all x in D. For any permutation function $f(.)$, let the function $F(.)$ of any vector $\mathbf{x} = \{x_1, .., x_k\}$ be defined as $F(\mathbf{x}) = \{f(x_1), ..., f(x_k)\}$. If strategies are symmetric an agent playing a_i after observing signal s_i and history $\boldsymbol{a}_{<i} = \{a_1, ..., a_{i-1}\}$ should play $f(a_i)$ after observing signal $f(s_i)$ and history $f(a_1), ..., f(a_{i-1})$:

Definition 1.1. *A pure strategy profile $\boldsymbol{\alpha}$ is **symmetric** if and only if*

$$\alpha_j[F(\boldsymbol{a}_{<j}), f(s_j)] = f[\alpha_j(\boldsymbol{a}_{<j}, s_j)] \tag{1.5}$$

for all permutation functions $f(.)$ and all j in all instantiations of the game $G\left(\rho; n, \{U(.)\}_{i=1}^{n}\right)$.

Attention is restricted to symmetric equilibria because it seems logical, given the uniformity of nature's choice, that agents should not have prior beliefs or strategies conditioned on particular elements in $[0,1]$.

The effect of i's choice on her expected utility can be decomposed into the expected private payoff to an action and the effect of that action on the social harm that i suffers. For a history of actions, $\boldsymbol{a}_{<j}$; an agent's signal s_j; and some strategy profile, $\boldsymbol{\alpha}$; i's expectation of the private payoff to some action a is denoted as

$$E[v(a)|\boldsymbol{a}_{<i}, s_i; \boldsymbol{\alpha}]. \tag{1.6}$$

Agents face a tradeoff between the private payoff of an action and that action's effect on the behaviour of subsequent agents. An action with a lower expected private payoff, c for example, may be preferred because it causes fewer agents to defect. An action that may cause more agents to defect, d^* for example, may be preferred because it has a higher private payoff. Agent i can only affect the social harm caused by future agents: $X(\boldsymbol{a}_{>i})$ where $\boldsymbol{a}_{>i} = \{a_j | j > i\}$. Hence, the optimal action of i is

$$a_i = \underset{a}{\operatorname{argmax}}\, E[v(a) - X(\boldsymbol{a}_{>i})|a, \boldsymbol{a}_{<i}, s_i; \boldsymbol{\alpha}]. \tag{1.7}$$

Before proceeding further it is useful to define two particular action profiles. If all agents prior to i have cooperated we say $\boldsymbol{a}_{<i} = \mathbf{c}$ and if all agents subsequent to i defect we say $\boldsymbol{a}_{>j} = \mathbf{d}^*$. Where $\mathbf{c} = \{c, ..., c\}$ and $\mathbf{d}^* = \{d^*, ..., d^*\}$ are vectors of arbitrary length.

Informed agents acting after histories in which no agent has defected have the largest potential impact on the number of agents defecting. This is because if an agent defects on path in a symmetric pure strategy equilibria then all subsequent agents defect. This is stated formally in the proposition below.

Proposition 1.1. *On the path of all symmetric pure strategy equilibria, if $a_i = d^*$ then all subsequent agents defect:*

$$a_i = d^* \Rightarrow \boldsymbol{a}_{>i} = \mathbf{d}^* = \{d^*, ..., d^*\}. \tag{1.8}$$

Proposition 1.1 implies that all agents defect after the very first defection. Hence, we only need the intuition for the case where i is the first agent to defect. Suppose then that no agent has yet defected. Clearly all agents know the exact action that i will take if uninformed.

Definition 1.2. *The symmetric pure strategy action of an uninformed agent is a_i^U:*

$$\alpha_i(\boldsymbol{a}_{<i}, \emptyset) = a_i^U. \tag{1.9}$$

Uninformed i does not know the value of d^* and so an uninformed i does not defect with probability one. It then follows that i defects only if i is informed and $\alpha_i(\boldsymbol{a}_{<i}, d^*) = d^*$ is i's pure strategy. Hence, if $a_i = d^*$ then all agents know that i is informed and $a_i = d^*$ because they can see that $a_i \neq a_i^U$. A simple backward induction argument then follows: the last agent, on path, will defect, as will the second to last, and so on back to the first agent subsequent to i. Once an agent defects in equilibrium then the knowledge of how to profit from doing bad has been irreversibly revealed and all agents defect.

This leads naturally to the importance of an agent's 'influence'. Consider an equilibrium in which agent i is the first agent to defect. Agent i can pretend to be uninformed and play the uninformed action instead of defecting. Proposition 1.1 states that all agents subsequent to i defect if i defects. Hence, pretending to be uninformed implies some agents, in expectation, do not defect when they would have if i had played her equilibrium strategy. The number of agents that she prevents from defecting is at least that group of uninformed agents acting after i but before the next informed agent. This group of uninformed agents, subsequent to i but prior to the next informed agent, whose behaviour is so affected by a_i, is defined below.[3]

[3] One key difference between the formal definition of an agent's influence and the common english meaning of 'influence' is that in this paper 'influence' refers to a group. Another key difference is that we

Definition 1.3. *The **influence** of agent i is the set of uninformed agents subsequent to i but before the next informed agent j': $\{i+1, \dots, j'\}$. The number of agents in i's influence is denoted by I_i.*

The expected size of i's influence, $E[I_i]$, will be critical in determining equilibrium play and has several important properties. The last agent, n, has no agents in her influence. The expected size of an agent i's influence will be determined exactly by the probability of independent learning, ρ, and the number of agents acting after i, $n-i$. $E[I_i]$ is monotonically increasing in $n-i$ and decreasing in ρ. Lemma 1.1 below defines $E[I_i]$ for any given i.

Lemma 1.1. *The expected size of i's influence is*

$$E[I_i] = \frac{(1-\rho) - (1-\rho)^{n+1-i}}{\rho}. \tag{1.10}$$

Note that the expected social harm that i suffers if all i's influence defects is then $\kappa E[I_i]$. I shall refer to $\kappa E[I_i]$ as "the potential harm of i's influence". We can now define a symmetric pure strategy equilibrium that exists for all instances of the anti-social learning game above.

Proposition 1.2. *For any $G(\rho; n, \{U(.)\}_{i=1}^n)$ a symmetric equilibrium exists where:*

$$\alpha_j(\boldsymbol{a}_{<j} = \mathbf{c}, s_j = \emptyset) = c. \tag{1.11}$$

$$\alpha_j(\boldsymbol{a}_{<j} = \mathbf{c}, s_j = d^*) = c, \qquad \textit{for } \kappa E[I_j] > \bar{v}. \tag{1.12}$$

$$\alpha_j(\boldsymbol{a}_{<j} = \mathbf{c}, s_j = d^*) = d^*, \qquad \textit{for } \kappa E[I_j] \leq \bar{v}. \tag{1.13}$$

$$\alpha_j(\boldsymbol{a}_{<j} \neq \mathbf{c}, s_j = d^*) = d^*. \tag{1.14}$$

$$\alpha_j(\boldsymbol{a}_{<j} \neq \mathbf{c}, s_j = \emptyset) = a_{j^*}, \qquad j^* = \max\{i < j | a_i \neq c\} \tag{1.15}$$

Uninformed j believes $v(a_{j^*}) = \bar{v}$ if $\boldsymbol{a}_{<j}$ is inconsistent with on path play. Informed j's beliefs about the payoffs to actions do not change off path.

On path, uninformed agents cooperate until some agent defects. Informed agent i cooperates if the potential harm of her influence, $\kappa E[I_i]$, is larger than the private payoff from defecting. The first agent to defect is the first informed agent where the potential harm of her influence is less than or equal to the private payoff from defecting. It follows immediately from the argument used to demonstrate proposition 1.1 that all agents subsequent to the first defecting agent then defect.

might argue, in some cases, that the number of agents influenced by an action a_i is larger or smaller than the number of agents in i's influence.

On and off path, an uninformed j believes that a_{j^*}, the most recently played action in D, has the highest private payoff. The informed always know the value of d^* and that it has the highest private payoff. There are two cases where agents are off path but do not know that they are off path. The first case is where an informed agent h cooperates after a cooperative history and $\kappa E[I_h] < \bar{v}$. The game is off path but no subsequent agent realises it because they do not know that h was informed. Hence, agents can only know that they are off path if some prior agent has not cooperated.

The second case is where some h such that $\kappa E[I_h] < \bar{v}$ blunders with $a_h \in D \backslash \{d^*\}$. In this case informed agents know that they are off path. Uninformed agents, on the other hand, will not know they are off path unless they are subsequent to some i who plays $a_i \neq a_h$. However, uninformed j still believes that a_{j^*} has the highest private payoff whether or not he realises that the game is off path. Consequently, the first agent i to know she is off path also knows that each subsequent agent will believe, for any action of hers, that some specific action in D maximises private payoffs.

These beliefs imply that if i knows she is off path then she also knows that all subsequent agents will play some action in D independently of her own action. To see this, consider that, as he has no influence, the last agent always maximises private payoffs. Consequently, he will always play $a_n = d^*$ if informed or $a_{n^*} \in D$ if uninformed. Hence, the social harm that the second to last agent suffers is independent of her own action. It is then optimal for the second to last agent to maximises her private payoffs by choosing $a_{n-1} = d^*$ if informed or $a_{n-1^*} \in D$ if uninformed. Exactly, the same logic then applies to the third to last agent, and so on, back to the first agent i that realises she is off path. It should be clear from the above arguments that, on and off path, if an agent doesn't cooperate after a cooperative history then all subsequent agents choose some action in D.

The intuition for the strategies of uninformed agents after a cooperative history follows immediately from the path of play that is implemented by an agent not cooperating. If an uninformed agent doesn't cooperate after a cooperative history they get a lower expected private payoff and, as all subsequent agents then choose an action in D, suffer a higher expected level of social harm. The same argument holds for an informed agent blundering, choosing an action in $D \backslash \{d^*\}$, after a cooperative history.

If an informed agent i defects, rather than cooperating, after a cooperative history she increases the number of agents defecting by at least the number of agents in her influence. If the potential harm of her influence is greater than the private payoff from defecting, $\kappa E[I_i] > \bar{v}$, then defecting causes her more expected social harm than she gains from the higher private payoff. Therefore, cooperating after a cooperative history is optimal when the potential harm of an informed agent's influence is larger than the private payoff of defecting.

Finally, we consider the optimality of defecting after a cooperative history when an informed agent's harmful influence is less than or equal to the private payoff from defect, $\kappa E[I_i] \leq \bar{v}$. The strategy of subsequent informed agents is to defect. Cooperating instead of defecting can then only reduce the number of agents defecting by the number of agents in i's influence. Hence, given the strategy of subsequent agents, it is optimal to defect as the potential harm of i's influence is less than or equal to the private payoff from defecting. It then remains to be shown that it is optimal for agents subsequent to i to defect if informed. The last agent always defects if informed and hence agents subsequent to $n-1$ defect if informed. Thus, $n-1$ defects if informed and $\kappa E[I_{n-1}] \leq \bar{v}$; so then does $n-2$ if $\kappa E[I_{n-2}] \leq \bar{v}$; and so on for all informed agents such that $\kappa E[I_i] \leq \bar{v}$.

The equilibrium defined in proposition 1.2 is not the only symmetric pure strategy equilibrium of the game. However, the play implemented by the equilibrium in proposition 1.2, for any particular instance of the anti-social learning game above, is identical for all symmetric pure strategy equilibria. Theorem 1.1 below characterises on path behaviour for all such equilibria.

Theorem 1.1. *On path play of all $i \in N$, in all symmetric pure strategy equilibria, can be characterised as follows:*

1) For uninformed i:

$$a_i = c \text{ if } \boldsymbol{a}_{<i} = \mathbf{c}, \tag{1.16}$$

$$a_i = d^* \text{ if } d^* \in \boldsymbol{a}_{<i}. \tag{1.17}$$

2) For informed i:

$$a_i = c \text{ if } \kappa E[I_i] > \bar{v}, \tag{1.18}$$

$$a_i \in \{c, d^*\} \text{ if } \kappa E[I_i] = \bar{v}. \tag{1.19}$$

$$a_i = d^* \text{ if } \kappa E[I_i] < \bar{v}. \tag{1.20}$$

This theorem implies that, on the path of all symmetric pure strategy equilibria, uninformed agents cooperate until some agent defects. Informed agent i cooperates if the potential harm of her influence, $\kappa E[I_i]$, is larger than the private payoff from defecting. The first agent to defect is the first informed agent where the potential harm of her influence is less than or equal to the private payoff from defecting. All agents subsequent to the first defecting agent defect. This is exactly the same behaviour as on the path of the equilibrium defined in proposition 1.2.

The intuition for this result is clearest if we just consider the choice between cooperate and defect. Also, suppose all agents act after a cooperative history or a history in which some agent has defected.[4] Cooperating after a cooperative history maximises uninformed agents expected private payoffs as they do not know how to defect; hence they cooperate. After an agent has defected then all agents defect, from proposition 1.1, so uninformed agents must defect after some other agent has defected.

An informed agent will not defect after a cooperative history if the potential harm of their influence is larger than the private payoff from defecting. Suppose it were an equilibrium for an informed agent i to defect after a cooperative history. If that agent pretends to be uninformed and cooperates then all the agents in her influence cooperate also. This reduces the social harm that she suffers by at least $\kappa E[I_i]$. If this is greater than the private payoff from defect it must then be profitable to deviate and cooperate. Hence, it cannot be an equilibrium strategy to defect when $\kappa E[I_i] > \bar{v}$ and so i cooperates when $\kappa E[I_i] > \bar{v}$.

If $\kappa E[I_i] < \bar{v}$ and informed agents' behaviour is as characterised in theorem 1.1 then the next informed agent and all subsequent agents defect. It then follows from proposition 1.1 that if i defects rather than cooperating she increases the number of agents defecting by exactly the number of agents in her influence. In which case defecting rather than cooperating increases her private payoff by $\bar{v}$ and the expected social harm that she suffers by $\kappa E[I_i]$. Thus, if $\kappa E[I_i] < \bar{v}$ and all subsequent informed agents defect then i defects.

The last agent always defects if informed. Therefore, it is the case that all agents subsequent to the second to last agent defect if informed. It then follows that the second to last agent defects if informed and $\kappa E[I_{n-1}] < \bar{v}$. In which case the third to last agent knows all subsequent agents defect if informed and chooses to defect if informed; and so on until an agent i such that $\kappa E[I_i] > \bar{v}$. This backward induction argument gives the intuition for agents acting after a cooperative history. For informed agents acting after a non-cooperative history in which some agent has already defected their behaviour follows from proposition 1.1. In the case that $\kappa E[I_i] = \bar{v}$ then i is indifferent between cooperating and defecting. This concludes the intuition for theorem 1.1.

[4] It is shown in the formal proof that agents never blunder, choose an action in $D \backslash \{d^*\}$, on the path of a symmetric pure strategy equilibrium.

1.3.2 Non-Symmetric Equilibria.

It is worth discussing, informally, the effect of relaxing symmetry. Relaxing symmetry allows for the existence of a strange set of equilibria where agents are forced by off-path beliefs to make "sacrificial" blunders in $[0, 1]$.

Consider a game where the potential harm of the first agent's influence is greater than $\kappa E[I_1] > (\bar{v} - \underline{v})$. In all symmetric equilibria this agent cooperates whether informed or uninformed. Suppose, however, a case where all uninformed agents believe that the action 0.5 is defect, $v(0.5) = \bar{v}$, if they are off the equilibrium path. Given these off path beliefs there is an equilibrium where the first agent always plays 0.5.

Such an equilibrium can have similar play to that in proposition 1.2. The first agent to defect on path being the first informed agent i such that $\kappa E[I_i] \leq \bar{v}$. All agents prior to the first defecting agent cooperate on path except for the very first agent in the game. The first agent plays $a_1 = 0.5$. However, off path all uninformed agents play 0.5 and all informed agents play d^*. The off path play follows from a backward induction argument and the off path beliefs of uninformed agents; they believe 0.5 maximises private payoffs. These off path beliefs and strategies imply that if the first agent does not play 0.5 all the uninformed agents will play 0.5 and all the informed agents will defect. This increases the expected social harm that the first agent suffers by at least $\kappa E[I_1] > (\bar{v} - \underline{v})$. Hence, it is optimal for the first agent to play 0.5 whether informed or uninformed.[5]

Note that the first agent playing 0.5 reveals no information about d^*. It is forced by non-symmetric off-path beliefs about the elements in $[0, 1]$. It needn't be the first agent who has to make the sacrificial blunder either. Any agent i can be forced by these kinds of non-symmetric off path beliefs to make a sacrificial blunder so long as $\kappa E[I_i] > (\bar{v} - \underline{v})$. Moreover, it needn't be just one agent who is forced to make a sacrificial blunder. There can be equilibria in which all agents for which $\kappa E[I_i] > (\bar{v} - \underline{v})$ are forced to make some sacrificial blunder. Leading to a long line of agents playing 0.5; then agents playing cooperate; then the first informed agent j^* such that $\kappa E[I_{j^*}] \leq \bar{v}$ defecting; and agents subsequent to j^* also defecting.

Such equilibria seem quite bizarre. It is not even obvious in what kind of situation we could even imagine them occurring. The attention of this paper was restricted to symmetric equilibria because all actions in D are ex-ante identical. It seemed unrealistic to suppose that agents should have ex-ante beliefs or strategies regarding particular elements in D. For example, it is unclear why agents should have any particular belief, ex-ante,

[5] The claims here should be obvious given the prior analysis in this paper. However, formal proofs are available from the author on request.

about the value 0.5 as it is identical, ex-ante, to all other elements in D. That these strange 'sacrificial' equilibria can occur when we move away from symmetry lends further justification to restricting beliefs and strategies so that they treat all elements in D as ex-ante identical.

1.4 The Effect of Social Learning on Cooperation

The probability of an individual being informed is the critical factor in determining whether social learning propagates or prevents uncooperative behaviour. If agents are unlikely to be informed then the informed agents expect to influence a large number of other actors. This causes the informed agents to internalise the social cost of their actions to a large extent and choose the socially optimal action. In which case social learning prevents uncooperative behaviour. Here, the possibility of social learning prevents anti-social behaviour and the possibility of anti-social behaviour prevents social learning.

An informed agent expects to have a less significant influence on the actions of others when others are more likely to be informed. If this influence is sufficiently small then all the informed agents choose to defect. This causes all the uninformed agents to defect also. In this case social learning propagates anti-social behaviour through the bad influence of the informed on the uninformed.

The effect of social learning on cooperation can be examined by comparing the behaviour in the dynamic game $G(\rho; n, \{U(.)\}_{i=1}^{n})$, above, to an equivalent static game $S(\rho; n, \{U(.)\}_{i=1}^{n})$. In the static game there is no social learning as agents do not observe the actions of other agents. The static game $S(\rho; n, \{U(.)\}_{i=1}^{n})$ has a dominant strategy equilibrium where all informed agents play d^* and all uninformed agents cooperate. In expectation the proportion of agents who defect in the static game $S(\rho; n, \{U(.)\}_{i=1}^{n})$ is ρ.

Theorem 1.2 below states that no informed agent cooperates in the dynamic game $G(\rho; n, \{U(.)\}_{i=1}^{n})$ if ρ is sufficiently high relative to a ratio of the social harm of defecting and the private payoff of d^*.

Theorem 1.2. *In any symmetric pure strategy equilibrium the first informed agent and all subsequent agents defect if:*

$$\rho \geq \bar{\rho} = \frac{\kappa}{\kappa + \bar{v}}. \tag{1.21}$$

Proof: From theorem 1.1 informed agent i and all subsequent agents defect in a symmetric pure strategy equilibrium if $\bar{v} > \kappa E[I_i]$. It can be shown that if $\rho > \bar{\rho}$ then $\bar{v} > \kappa E[I_i]$ for all agents in any finite population.

The expected influence of i is

$$E[I_i] = \frac{(1-\rho)-(1-\rho)^{n+1-i}}{\rho} < \frac{1-\rho}{\rho} \text{ for all } i \geq 1 \text{ and all } n \in \mathbb{N}.$$

If $\rho \geq \bar{\rho}$ then

$$\bar{v} \geq \kappa \frac{1-\rho}{\rho},$$

and therefore,

$$\bar{v} > \kappa E[I_i] \text{ for all } i \geq 1 \text{ and all } n \in \mathbb{N}.$$

□

It follows immediately from theorem 1.2 that social learning increases the number of agents who defect when ρ and n are sufficiently large. Hence, corollary 1.3 below gives sufficient conditions under which social learning causes a deterioration in cooperative behaviour.

Corollary 1.3. *The expected number of agents defecting in any symmetric pure strategy equilibrium of the dynamic game* $G(\rho; n, \{U(.)\}_{i=1}^{n})$ *is greater than in the static game* $S(\rho; n, \{U(.)\}_{i=1}^{n})$ *if* $\rho \geq \bar{\rho}$ *and* $n > \frac{1}{\rho}$

Proof: In $G(\rho; n, \{U(.)\}_{i=1}^{n})$, for $\rho \geq \bar{\rho}$, the expected number of agents cooperating in the dynamic game $G(\rho; n, \{U(.)\}_{i=1}^{n})$ is bounded above by

$$\frac{1-\rho}{\rho}. \tag{1.22}$$

Hence, the expected number of agents defecting in $G(\rho; n, \{U(.)\}_{i=1}^{n})$ is bounded below by

$$n - \frac{1-\rho}{\rho}. \tag{1.23}$$

The expected number of agents defecting in $S(\rho; n, \{U(.)\}_{i=1}^{n})$ is ρn. Therefore, the expected number of agents defecting in the dynamic game less the expected number of agents defecting in the static game is

$$n - \frac{1-\rho}{\rho} - \rho n = (1-\rho)n - \frac{1-\rho}{\rho}. \tag{1.24}$$

The right hand side of (1.24) divided by $1-\rho$ is

$$n - \frac{1}{\rho} > 0. \tag{1.25}$$

□

Theorem 1.4 below states that the population is almost entirely cooperative for very large populations when $\rho < \bar{\rho}$.

Theorem 1.4. *In any symmetric pure strategy equilibrium of the game* $G(\rho; n, \{U(.)\}_{i=1}^{n})$ *with* $\rho < \bar{\rho}$ *the proportion of agents cooperating tends to one as the population size tends to infinity.*

Proof:

Take a fixed set of parameters ρ, κ and $\bar{v}$ such that $\rho < \bar{\rho}$. The expected harm from the first agent's influence choosing d^* is

$$\kappa E[I_1] = \kappa \frac{(1-\rho) - (1-\rho)^n}{\rho}. \tag{1.26}$$

Because $\rho < \bar{\rho}$, equation (1.26) tends to

$$\kappa \frac{(1-\rho)}{\rho} > \bar{v}$$

as n tends to infinity. Therefore, there exists some finite $n = \underline{n}$ such that $\kappa E[I_1] > \bar{v}$ and $\kappa E[I_i] < \bar{v}$ for all $i > 1$. $\underline{n}$ is any n that satisfies

$$\kappa \frac{(1-\rho) - (1-\rho)^{\underline{n}}}{\rho} > \bar{v} > \kappa \frac{(1-\rho) - (1-\rho)^{\underline{n}-1}}{\rho}. \tag{1.27}$$

Suppose a population $n' > \underline{n}$. It follows that the expected social harm from i's influence playing d^* is greater than $\bar{v}$ if and only if there are fewer than $\underline{n}-1$ agents acting subsequent to i. That is

$$\kappa E[I_i] = \kappa \frac{(1-\rho) - (1-\rho)^{n'+1-i}}{\rho} < \bar{v} \text{ if and only if } n' - i < \underline{n} - 1. \tag{1.28}$$

Hence, the number of agents that defect in an informative equilibrium where $\rho < \bar{\rho}$ is bounded above by the finite integer $\underline{n}-1$. The proportion of agents cooperating for $n' > \underline{n}$ is then bounded below by

$$\frac{n' + 1 - \underline{n}}{n'}$$

which tends to one as n' tends to infinity.

□

From theorem 1.4 it can easily be shown that social learning decreases the number of agents who defect when $\rho < \bar{\rho}$ and n is sufficiently large.

Corollary 1.5. *There always exists a finite* $\bar{n}$ *such that the expected number of agents defecting in any symmetric pure strategy equilibrium of the dynamic game* $G(\rho; n, \{U(.)\}_{i=1}^{n})$ *is less than in the static game* $S(\rho; n, \{U(.)\}_{i=1}^{n})$ *if* $\rho < \bar{\rho}$ *and* $n > \bar{n}$.

Proof: In the static game $S(\rho; n, \{U(.)\}_{i=1}^{n})$ the expected number of agents defecting is ρn. From theorem 1.4 the proportion of agents defecting in the dynamic game $G(\rho; n, \{U(.)\}_{i=1}^{n})$ tends to zero as n tends to infinity. Therefore, there exists some finite $\bar{n}$ such that if $n > \bar{n}$ the expected number of agents defecting is less than $\rho n > 0$.

□

In this section it has been shown that when the probability of being informed is greater than $\bar{\rho}$ that social learning propagates uncooperative behaviour through the bad influence of the informed on the uninformed. However, when $\rho < \bar{\rho}$, and the population is large, then social learning prevents uncooperative behaviour by causing the informed to internalise the social cost of their actions due to their potential bad influence on the uninformed. Consequently, social learning can reduce uncooperative behaviour in populations when the probability of independent learning is low.

1.5 Conclusion

The anti-social learning game in this paper provides an insight into the potential importance of social learning for norms of cooperation. This insight is that the probability of independent learning can determine whether social learning propagates uncooperative behaviour or prevents it. When people are likely to learn in an independent fashion then social learning tends to propagate uncooperative behaviour. When people are unlikely to learn in an independent fashion then social learning can prevent people from being uncooperative.

It is easy to observe cases when uncooperative behaviour is propagated via social learning. It is, however, quite difficult to observe cases where the possibility of social learning prevents uncooperative behaviour. To see that uncooperative behaviour is being prevented we would need to know that the socially harmful activity exists; that it is privately profitable for some agents; and that some of the economic actors have this information too.

To illustrate this point, consider the lake fishermen discussed in the introduction. Recall that there was one lake fisherman with knowledge of blast fishing. This fisherman did not adopt the practice because he did not want other fishermen on his lake to adopt the practice. None of the other lake fishermen knew about blast fishing and would not have realised that the potential for social learning had prevented some fisherman from adopting this harmful practice.

In many cases we are like these uninformed lake fishermen. We do not know an uncooperative activity is not happening because we do not know such an activity is possible, or profitable. In other cases we may be aware of a profitable uncooperative activity but not know that the relevant actors have this information. For example, we might know of blast fishing but not know that some of the lake fishermen know of blast fishing. In which case the absence of blast fishing may only indicate that no fisherman on the lake is aware of blast fishing.

We then only see that the potential for social learning prevents informed fishermen from practising blast fishing when we both know about the practice of blast fishing *and* that some fishermen know of the practice. In general we can only have concrete evidence for this preventative role of social learning when we are privy to both the information of the 'informed' agents and we know that some agents are informed. It may well be difficult for economists to acquire such information while the uncooperative behaviour is still being repressed.

It may, however, be fruitful to examine cases where a new form of uncooperative behaviour has taken off. It would be partial evidence for the suppression of uncooperative behaviour if a surge in new forms of uncooperative behaviour had been preceded by large increases in the proportion of people, or organisations, which we would characterise as having a high probability of independent learning. This might be an increase in the proportion of people who are intelligent; possess expertise; or are well connected outside the group. In the case where organisations are the unit of analysis this might be an increase in the proportion of organisations which have high expenditure on research and development or strategic analysis.

An important implication then of this model is that organisations may want to restrict the proportion of people who are intelligent; innovative; or have special expertise. This is a possible factor contributing to a desire of firms to not employ 'over-qualified' workers. Organisations or groups may also like to make experimentation highly costly in order to stop people from independently discovering socially harmful activities. This might be a contributing factor for the many taboos and high demands for conformity demonstrated in many tribal societies. In the same vein this may help explain why some groups or societies choose to be insular, so they can stop their members from learning harmful practices from outsiders.

Another implication of this model is that it is not necessarily a good idea to isolate people with ‘bad behaviour’ from people with ‘good behaviour’. The fear is that those with bad behaviour will be a bad influence on those with good behaviour. However, by putting the badly behaved in with the well behaved we increase the number of people who can be influenced by an individuals bad behaviour. If the numbers of the badly behaved are sufficiently small relative to the well behaved then the model in this paper predicts that they will suppress their own bad behaviour. Their new found influence, or responsibility, causes them to internalise the social cost of their actions and turn over a new leaf.

Appendix

PROOF of Lemma 1.1:

The probability of there being exactly $j \leq n-i$ agents in i's influence is the probability of there being at least j less the probability of there being at least $j+1$:

$$(1-\rho)^j - (1-\rho)^{j+1}.$$

Hence, the expected number of agents in an informed agent's influence is given by:

$$E[I_i] = \sum_{j=1}^{n-i} j[(1-\rho)^j - (1-\rho)^{j+1}], \tag{1.29}$$

which immediately simplifies to give

$$E[I_i] = \sum_{j=1}^{n-i}(1-\rho)^j = \frac{(1-\rho)-(1-\rho)^{n+1-i}}{\rho}. \tag{1.30}$$

□

PROOF of Proposition 1.1:

(i) If $a_i = d^*$ on path in a symmetric pure strategy equilibrium then it is common knowledge for all $j \geq i$ that $a_i = d^*$ with probability one.

All informed agents know d^* and so we only need consider whether uninformed agents can infer that $a_i = d^*$. Uninformed j observes $a_i \in D$ and not the informational status of i, s_i. Let $a_j^U = \alpha_i(\boldsymbol{a}_{<i}, \emptyset)$ and $a_i^I = \alpha_i(\boldsymbol{a}_{<i}, d^*)$ be the on path actions of an uninformed i and an informed i respectively.

Consider first the case where $d^* \notin \boldsymbol{a}_{<i}$ and $a_i^I = d^*$. In this case $a_i^U \neq a_i^I$ with probability 1 as an uninformed i does not know the value of d^* and a_i^U could then only equal d^* by

chance, i.e. with probability zero. All $j > i$ know the exact value of a_i^U and that $a_i^I = d^*$. Therefore, if $a_i \neq a_i^U$ then all $j > i$ know that i is informed and that $a_i = d^*$.

This proves *(i)* for the case where $d^* \notin \boldsymbol{a}_{<i}$ and $a_i^I = d^*$. The case where $d^* \notin \boldsymbol{a}_{<i}$ and $a_i^U = d^*$ is a zero probability event. For the case where $d^* \in \boldsymbol{a}_{<i}$ let $h < i$ be the first agent to defect, $a_h = d^*$. From the above analysis it is common knowledge that $a_h = d^*$. $a_i = d^*$ only if $a_i = a_h$ and hence it is common knowledge amongst all $j > i$ that $a_i = d^*$. This concludes the proof of *(i)*.

(ii) If it is common knowledge for all $j > i$ on path that $a_i = d^*$ then all subsequent agents defect on path.

Let it be common knowledge that $a_i = d^*$. Suppose $\alpha(\boldsymbol{a}_{<k}, s_k) = d^*$ for all $k > j > i$ and $a_j \neq d^*$ on path. j knows that $a_i = d^*$ whether j is informed or uninformed. If j deviates and plays $a_j = d^*$ then j has a higher private payoff and cannot increase the social harm that she suffers. $a_j = d^*$ would then be a profitable deviation. Therefore it cannot be the case that $\alpha(\boldsymbol{a}_{<k}, s_k) = d^*$ for all $k > j > i$ and $a_j \neq d^*$ on path. It follows that,

$$\alpha(\boldsymbol{a}_{<k}, s_k) = d^* \text{ on path } \forall k > j \Rightarrow a_j = d^* \text{ on path,} \tag{1.31}$$

and thus

$$\alpha(\boldsymbol{a}_{<k}, s_k) = d^* \text{ on path } \forall k > j \Rightarrow \alpha(\boldsymbol{a}_{<k}, s_k) = d^* \text{ on path } \forall k > j - 1. \tag{1.32}$$

in all equilibria where it is common knowledge that $a_i = d^*$.

The last agent, n, maximises private payoffs and so

$$\alpha(\boldsymbol{a}_{<n}, s_n) = d^* \text{ if } n > i. \tag{1.33}$$

Given (1.33) we can iteratively apply (1.32) to get *(ii)* by backwards induction.

Proposition 1.1 then follows immediately from *(i)* and *(ii)*. □

PROOF of Proposition 1.2:

First it is shown that the two sub-strategies that define the response to noncooperative histories, (1.14) and (1.15), are both optimal on and off the equilibrium path. Sub-strategy (1.14), $\alpha_j(\boldsymbol{a}_{<j} \neq \mathbf{c}, s_j = d^*) = d^*$, defines the behaviour of informed agents. Sub-strategy (1.15), $\alpha_j(\boldsymbol{a}_{<j} \neq \mathbf{c}, s_j = \emptyset) = a_{j^*}$ where $j^* = \max\{i < j | a_i \neq c\}$, defines the behaviour of uninformed agents.

Consider an agent acting after a noncooperative history where all subsequent agents will choose some action in D after any noncooperative history. It then follows that this agent cannot affect the externalities generated by subsequent agents:

$$[\mathbf{a_{<j}} \neq \mathbf{c} \text{ and } \alpha_k(\boldsymbol{a}_{<k} \neq \mathbf{c}, s_k) \in D \ \forall k > j] \Rightarrow X(\boldsymbol{a}_{>j}) \perp\!\!\!\perp a_j. \tag{1.34}$$

If an agent's action does not affect the externalities generated by subsequent agents then the agent maximises expected private payoff:

$$X(\boldsymbol{a}_{>j}) \perp\!\!\!\perp a_j \Rightarrow \underset{a_j}{\operatorname{argmax}}\, E[v(a_j)|\mathbf{a_{<j}}, s_j; \boldsymbol{\alpha}] = \underset{a}{\operatorname{argmax}}\, E[U(a_j, \boldsymbol{a}_{-j}, d^*)|\mathbf{a_{<j}}, s_j; \boldsymbol{\alpha}]. \tag{1.35}$$

Given equilibrium behaviour and off path beliefs, after some $i < j$ has played $a_i \in D$ uninformed agent j believes a_{j^*} maximises private payoffs. An informed agent knows the value of d^*, on and off path, and that it maximises private payoffs. Therefore, an action in D maximises the private payoffs of uninformed and informed agents after a noncooperative history:

$$\underset{a_j}{\operatorname{argmax}}\, E[v(a_j)|\mathbf{a_{<j}} \neq \mathbf{c}, s_j; \boldsymbol{\alpha}] \in \{a_{n^*}, d^*\} \in D. \tag{1.36}$$

It then follows immediately from (1.34), (1.35) and (1.36) that j plays an action in D after a noncooperative history if all subsequent agents have a strategy of choosing some action in D after any noncooperative history:

$$\alpha_k(\boldsymbol{a}_{<k} \neq \mathbf{c}, s_k) \in D \ \forall k > j \Rightarrow \alpha_k(\boldsymbol{a}_{<k} \neq \mathbf{c}, s_k) \in D \ \forall k > j-1 \tag{1.37}$$

The last agent, n, cannot affect the social harm generated by others. Hence, n maximises expected private payoff:

$$\alpha_n(\boldsymbol{a}_{<n} \neq \mathbf{c}, s_n) \in \{a_{n^*}, d^*\} \in D. \tag{1.38}$$

Consequently,

$$\alpha_k(\boldsymbol{a}_{<k} \neq \mathbf{c}, s_k) \in D \quad \forall k > n-1. \tag{1.39}$$

Given (1.39) we can iteratively apply (1.37) to demonstrate the optimality of sub-strategies (1.14) and (1.15) by backwards induction.

Next we examine the sub-strategies that define the response to cooperative histories, (1.11), (1.12) and (1.13). Agent j always believes she is on the equilibrium path if $a_{<j} = \mathbf{c}$. Therefore, we only need to prove that (1.11), (1.12) and (1.13) are optimal for agent j who believes she is on path.

It has just been shown that if $a_i \in D$ then $a_j \in D$ for all $j > i$, on and off the equilibrium path. Hence, agent i plays $a_i \in D$ only if a_i maximises i's expected private payoff. That is

$$[\boldsymbol{a}_{<i} = \mathbf{c} \text{ and } a_i \in D] \Rightarrow a_i = \underset{a}{\operatorname{argmax}} \, E[v(a)|\boldsymbol{a}_{<i}, s_i; \alpha]. \tag{1.40}$$

Defecting does not maximises expected private payoff for uninformed i if $\boldsymbol{a}_{<i} = \mathbf{c}$. So it follows from (1.40) that the strategy defined in (1.11), $\alpha_j(\boldsymbol{a}_{<j} = \mathbf{c}, s_j = \emptyset) = c$, is optimal.

Next, we examine the optimality of sub-strategies (1.12), $\alpha_j(\boldsymbol{a}_{<j} = \mathbf{c}, s_j = d^*) = c$ if $\kappa E[I_j] > \bar{v}$; and (1.13), $\alpha_j(\boldsymbol{a}_{<j} = \mathbf{c}, s_j = d^*) = d^*$ if $\kappa E[I_j] < \bar{v}$. From (1.40), if $\boldsymbol{a}_{<j} = \mathbf{c}$ then $a_j \notin D\backslash\{d\}$ as $a_j \in D\backslash\{d\}$ does not maximises expected private payoffs. Hence, we need only compare the expected utility from cooperate and defect to show optimality of the two sub-strategies:

$$E[U(a_j = d^*, \boldsymbol{a}_{-j}, d^*) - U(a_j = c, \boldsymbol{a}_{-j}, d^*)|\mathbf{a}_{<\mathbf{j}}, s_j = d^*; \boldsymbol{\alpha}]. \tag{1.41}$$

If expression (1.41) is less than zero when $\kappa E[I_j] > \bar{v}$ then sub-strategy (1.12) is optimal. Likewise, sub-strategy (1.13) is optimal if expression (1.41) is greater than or equal to zero when $\kappa E[I_j] \leq \bar{v}$.

From the strategies defined in (1.15) and (1.11) it follows that all j's influence plays d^* if $a_j = d^*$ and all j's influence play c if $a_j = c$ when $\boldsymbol{a}_{<j} = \mathbf{c}$. This implies the expected social harm suffered by j is greater by at least $\kappa E[I_j]$ for $a_j = d^*$ than for $a_j = c$. Therefore, $a_j = c$ is optimal if $\kappa E[I_j] > v(d^*) - v(c) = \bar{v}$; proving the optimality of sub-strategy (1.12).

If the strategy of agents subsequent to j is $\alpha_k(\boldsymbol{a}_{<k} = \mathbf{c}, s_k = d^*) = d^*$ for all $k > j$ then the expected social harm suffered by j is greater by exactly $\kappa E[I_j]$ for $a_j = d^*$ than for $a_j = c$. Hence

$$[\kappa E[I_j] > \bar{v} \text{ and } \alpha_k(\boldsymbol{a}_{<k} = \mathbf{c}, s_k = d^*) = d^*, \ \forall k > j] \Rightarrow \alpha_j(\boldsymbol{a}_{<j} = \mathbf{c}, s_j = d^*) = d^*. \tag{1.42}$$

The last agent maximises expected private payoffs so $\alpha_n(\boldsymbol{a}_{<n} = \mathbf{c}, s_n = d^*) = d^*$. Which implies that $\alpha_k(\boldsymbol{a}_{<k} = \mathbf{c}, s_k = d^*) = d^*$ for all $k > j = n - 1$. By backwards induction

$\alpha_j(\boldsymbol{a}_{<j} = \mathbf{c}, s_j = d^*) = d^*$ is optimal for all j such that $\kappa E[I_j] \leq v(d^*) - v(c) = \bar{v}$; proving the optimality of sub-strategy (1.13).

Finally, we prove symmetry. We can write sub-strategy (1.15) as $\alpha_j(\boldsymbol{a}_{<j} \neq \mathbf{c}, s_j = \emptyset) = \max\{a_i \in \boldsymbol{a}_{<j} | a_i \neq c\}$ and we immediately see that (1.15) satisfies symmetry from the fact that $f(\max\{a_i \in \boldsymbol{a}_{<j} | a_i \neq c\}) = \max\{a_i \in F(\boldsymbol{a}_{<j}) | a_i \neq c\}$ for all permutation functions f. We can write sub-strategy (1.14) as $\alpha_j(\boldsymbol{a}_{<j} \neq \mathbf{c}, s_j = d^*) = s_j$. Hence, $\alpha_j(F(\boldsymbol{a}_{<j}) \neq \mathbf{c}, f(s_j = d^*)) = f(s_j) = f(\alpha_j(\boldsymbol{a}_{<j} \neq \mathbf{c}, s_j = d^*)$. We can write sub-strategy (1.13) as $\alpha_j(\boldsymbol{a}_{<j} = \mathbf{c}, s_j = d^*) = s_j$ when $\kappa E[I_j] < \bar{v}$ and then show symmetry by the same argument as for sub-strategy (1.14). The symmetry of (1.11) and (1.12) follows immediately from the fact that $f(c) = c$.

$\square$

PROOF of Theorem 1.1:

First, it is shown that there is no history $\boldsymbol{a}_{<i}$ on path where $d^* \notin \boldsymbol{a}_{<i}$ and $\boldsymbol{a}_{<i} \neq \mathbf{c}$. This holds if no i plays $a_i = x \in D \backslash \{d^*\}$ after a cooperative history. For uninformed i, if $\alpha_i(\mathbf{c}, \emptyset) = x \neq c$ then $\alpha_i(F(\mathbf{c}), f(\emptyset)) = \alpha_i(\mathbf{c}, \emptyset) = x \neq f(\alpha_i(\mathbf{c}, \emptyset))$ for some permutation function $f(.)$. For informed i, suppose a permutation function such that $f(d^*) = d^*$ and $f(x) \neq x$. It is then obvious that the strategy $\alpha_i(\mathbf{c}, d^*) = x$ is not symmetric as $\alpha_i(F(\mathbf{c}), f(d^*)) = \alpha_i(\mathbf{c}, d^*) = x \neq f(x)$. Hence, agents only act after cooperative histories or histories containing d^*, i.e. on path it is the case that $d^* \in \boldsymbol{a}_{<i}$ or that $\boldsymbol{a}_{<i} = \mathbf{c}$.

The behaviour of uninformed agent i after a cooperative history must be c to satisfy symmetry, as shown above. It follows from proposition 1.1 that uninformed agent i plays $a_i = d^*$ if $d^* \in \boldsymbol{a}_{<i}$. This proves that uninformed agents on path behaviour is fully characterised by theorem 1.1.

Next, we see that on path behaviour of informed agent i can be completely characterised by $\kappa E[I_i]$ relative to $\bar{v}$. We begin with the behaviour of informed agents acting after cooperative histories. After a cooperative history, as shown above, an informed agent plays $a_i \in \{c, d^*\}$.

If $\alpha_i(\mathbf{c}, d^*) = d^*$ and i deviates by playing cooperate then no subsequent agent realises that they are off path and, given uninformed on path play, all agents in i's influence will cooperate. If i does defect then, from proposition 1.1, we know that all subsequent agents defect. Thus, by cooperating i reduces the expected social harm that i suffers by at least $\kappa E[I_i]$ and reduces private payoffs by $\bar{v}$. Therefore, $a_i = c$ for all informed i on path after a cooperative history if $\kappa E[I_i] > \bar{v}$.

This implies that all agents for which $\kappa E[I_i] \geq \bar{v}$ act after a cooperative history. Thus, the behaviour in (1.18) fully characterises the on path behaviour of informed agents for which $\kappa E[I_i] > \bar{v}$. It also follows immediately that the behaviour in (1.19) fully characterises the on path behaviour of informed agents for which $\kappa E[I_i] = \bar{v}$ as informed i plays $a_i \in \{c, d^*\}$ after a cooperative history.

Consider now an informed agent i where $\kappa E[I_i] < \bar{v}$ and $\boldsymbol{a}_{<i} = \mathbf{c}$. Suppose all subsequent informed agents have the following strategy:

$$\alpha_i(\mathbf{c}, d^*) = d^* \;\; \forall j > i. \tag{1.43}$$

The first informed agent after i (after i's influence) plays d^* and so then do all subsequent agents. If $a_i = d^*$ then the worst outcome is if all subsequent agents defect. If $a_i = c$ (which is consistent with on path play as others do not observe if i is informed) then the best outcome for i, given the strategy of subsequent informed agents, is that i's influence cooperates and then all subsequent agents defect. This implies that playing $a_i = d^*$ increases the expected social harm that i suffers by $\kappa E[I_i]$ at most. This is more than compensated for by the increase in the private payoff of i as $\kappa E[I_i] < \bar{v}$. Therefore, $a_i = d^*$ for informed i on path after a cooperative history where subsequent informed agents have the strategy defined in (1.43). Which implies the following for i such that $\kappa E[I_i] < \bar{v}$:

$$\alpha_i(\mathbf{c}, d^*) = d^* \;\; \forall j > i \Rightarrow \alpha_i(\mathbf{c}, d^*) = d^* \;\; \forall j > i - 1. \tag{1.44}$$

The last agent n maximises private payoffs and so $\alpha_n(\mathbf{c}, d^*) = d^*$. Iteratively applying (1.44) implies that $\alpha_i(\mathbf{c}, d^*) = d^*$ for all informed i such that $\kappa E[I_i] < \bar{v}$. From proposition 1.1 we know that all informed i play d^* on path if $d^* \in \boldsymbol{a}_{<i}$. As there is no history where $\boldsymbol{a}_{<i} \neq \mathbf{c}$ and $d^* \notin \boldsymbol{a}_{<i}$ this implies that (1.20) fully characterises the on path behaviour for informed i such that $\kappa E[I_i] > \bar{v}$.

□

Acknowledgements

I am grateful to Ed Hopkins, Jozsef Sakovics, Jakub Steiner, Keshav Dogra, Tatiana Kornienko, Stephen Morris, Stephen Durlauf, Philip Reny, Kohei Kawamura, Jonathan Thomas, David Pugh, Sean Brocklebank, Nick Vikander, and Jose V. Rodriguez Mora for their comments and criticisms on earlier draughts of this paper. This work was produced as a post-graduate student at the University of Edinburgh and as a visiting student at the University of Chicago. I received funding from the British Economic Social Research Council postgraduate studentship funding scheme and the Scottish Institute for Research in Economics.

Chapter 2

Self-Fulfilling Price Cycles

James Best and John Moore

This chapter presents a model of a self-fulfilling price cycle in an asset market. In this model the dividend stream of the economy's asset stock is constant yet price oscillates deterministically even though the underlying environment is stationary. This creates a model in which there is rational excess volatility - 'excess' in the sense that it does not reflect changes in dividend streams and 'rational' in that all agents are acting on their best information. The mechanism that we uncover is driven by endogenous variation in the investment horizons of the different market participants, informed and uninformed.

On even days, the price is high; on odd days it is low.

On even days, informed traders are willing to jettison their good assets, knowing that they can buy them back the next day, when the price is low. The anticipated drop in price more than offsets any potential loss in dividend. Because of these asset sales, the informed build up their cash holdings. Understanding that the market is flooded with good assets, the uninformed traders are willing to pay a high price. But their investment horizon is longer than that of the informed traders: their intention is to hold the assets they purchase, not to resell.

On odd days, the price is low because the uninformed recognise that the informed are using their cash holdings to cherry-pick good assets from the *market. Now the uninformed, like the informed, are investing short-term. Rather than buy-and-hold as they do with assets purchased on even days, on odd days the uninformed are buying to sell.*

Notice that, at the root of the model, there lies a credit constraint. Although the informed are flush with cash on odd days, they are not deep pockets. On each cherry that they pick out of the market, they earn a high return: buying cheap, selling dear. However they don't have enough cash to strip the market of cherries and thereby bid the price up.

2.1 Introduction

Shiller (1980) argues that the volatility in asset prices is not justified by subsequent changes in dividend streams. Nor, he argues, are they justified by changes in public information about fundamentals (Shiller, 1992). Such excess volatility is *prima facie* evidence that market participants are behaving irrationally. We ask if it is possible that such excess volatility be generated by purely rational market participants. Moreover, can it be explained from within the model – that is to say, endogenously, without exogenous variation?

It is known that bubbles can be part of a *stochastic* equilibrium (Abreu and Brunnermeier, 2003). In this paper we set ourselves the challenge of generating *deterministic* price movement in a perfectly stationary environment where asset prices reflect all publicly available information. In this economy the dividend stream of the asset stock would be constant while the asset prices would be volatile - a most extreme form of excess volatility. We show that it is possible.

In our model the dynamic interaction of adverse selection (we have informed and uninformed traders) and credit constraints (the informed traders do not have deep pockets) can endogenously generate variation in the investment horizons of the different market participants. With these simple ingredients, we demonstrate a saw-tooth equilibrium in which prices deterministically rise one day and fall the next, *ad infinitum*. On even days, the price is high; on odd days it is low.

On even days, informed traders are willing to jettison their good assets, knowing that they can buy them back the next day, when the price is low. The anticipated drop in price more than offsets any potential loss in dividend. Because of these asset sales, the informed build up their cash holdings. Understanding that the market is flooded with good assets, the uninformed traders are willing to pay a high price. But their investment horizon is longer than that of the informed traders: their intention is to hold the assets they purchase, not to resell.

On odd days, the price is low because the uninformed recognise that the informed are using their cash holdings to cherry-pick good assets from the market. Now the uninformed, like the informed, are investing short-term. Rather than buy-and-hold as they do with assets purchased on even days, on odd days the uninformed are buying to sell.

Notice that, at the root of the model, there lies a credit constraint. Although the informed are flush with cash on odd days, they are not deep pockets. On each cherry that they pick out of the market, they earn a high return: buying cheap, selling dear. However they don't have enough cash to strip the market of cherries and thereby bid the price up.

This paper sits within the literature on adverse selection that started with Akerlof (1970). Our model is an infinite horizon adverse selection model. For earlier examples of such models one can see Hendel and Lizzeri (1999) and Hendel et al. (2005) on dynamic adverse selection in the context of durable goods markets. These papers, as opposed to ours, do not look to examine or generate price volatility through this mechanism.

Dow (2004) develops an adverse selection model that examines the potential for self-fulfilling liquidity. In this paper there are two equilibria: a high liquidity equilibrium with narrow bid-ask spreads and a low proportion of trade being informational; and a low liquidity equilibrium with wide bid-ask spreads and a high proportion of trade being informational. Unlike the model in Dow (2004) we examine prices, as opposed to liquidity. We also examine adverse selection in a dynamic (infinite-horizon) setting as opposed to a one shot game. As a result the meaning of 'self-fulfilling' in the two papers is different. In Dow (2004) liquidity is self-fulfilling in the multiple equilibrium sense. There are two possible levels of liquidity and either level, high or low, justifies the equilibrium behaviour that justifies the level of liquidity. In our paper the price cycle is not self-fulfilling in the multiple equilibrium sense but in the sense that future prices justify the present price and the present price justifies future prices.

Eisfeldt (2004) and Guerrieri and Shimer (2013) use adverse selection to model liquidity in a dynamic infinite horizon setting. Aggregate dynamics within their models result from aggregate shocks. A key difference, then, between these papers and our own is that the aggregate dynamics in our model arise completely endogenously, in a perfectly stationary environment.

In Section II, we introduce the formal model. In Section III, we examine a constant price equilibrium. In section IV, we construct a self-fulfilling price cycle equilibrium. In section V, we provide a numerical example.

2.2 Model

The economy is discrete time, with a durable generic consumption good, fruit, and a single kind of asset, trees. At the start of each day trees mature with probability μ. The maturation process is *i.i.d.* across trees, age and time. When trees mature they bear fruit and immediately die. Trees that do not mature do not bear fruit and do not die.

There are two types of tree: high quality "cherry" trees and low quality "lemon" trees. A fraction γ of trees are cherry. When trees mature, lemon trees bear $l > 0$ (generic) fruit and cherry trees bear $h = l + \Delta$ (generic) fruit, where $\Delta > 0$.

There is a continuum of risk-neutral traders. At the end of each day traders die with probability δ. Agents learn earlier in the day, after trees have matured, if they are to die at the end of the day. Death is *i.i.d.* across traders, age and time. Traders consume fruit, only once, just before they die.

There are two types of trader: "informed" traders and "uninformed" traders. Informed traders can determine whether an unmatured tree is a cherry or a lemon; uninformed traders cannot. A fraction θ of traders are informed.

Each day – after the maturation of a fraction μ of the existing trees, but before a fraction δ of traders learn if they will die at the end of the day – a mass α of new traders are born endowed with a mass τ of new, unmatured, trees. A fraction γ of these new trees are cherry and a fraction θ of these new traders are informed. The overnight steady-state stock T of trees (of which γT are cherry) solves $\mu T = \tau$. And the overnight steady-state mass A of traders (of whom θA are informed) solves $\delta A = \alpha$.

There is a daily competitive spot market in which trees are traded for fruit. The market occurs after traders learn if they will die at the end of the day. That is, the timing on each day is as follows:

1. A fraction μ of the existing trees mature: they yield fruit and die.
2. A mass α of new traders (of whom a fraction θ are informed) are born endowed with a mass τ of new, unmatured, trees (of which a fraction γ are cherry).
3. A fraction δ of traders – including a fraction δ of the newly-born[1] – learn that they are dying today (see 5 below).
4. A market occurs in which trees are traded for fruit.
5. The dying traders consume their fruit and die.

We can save on notation by assuming $\tau = \mu$ and $\alpha = \delta$, so that $T = 1$ and $A = 1$.

This model typically has several equilibria. Given the stationary nature of the environment, it is unsurprising that there is an equilibrium where price is constant through time; see section 2 below.

We do not try to characterise all the non-constant-price equilibria. However, of particular interest to us is an equilibrium where the price deterministically oscillates: a high price on (e.g.) even days and a low price on odd days. We analyse this "saw-tooth" equilibrium in section 4. It is to this surprising equilibrium that we wish to draw the reader's attention.

[1]These traders live short lives.

There are features that are common to all the equilibria we study – i.e., common to a constant-price equilibrium and a saw-tooth equilibrium. Crucially, we assume that informed traders are able to cherry pick in the market – they only purchase cherry trees, to the extent that their budgets allow. Uninformed traders purchase the residual supply of trees. Of the trees purchased by the uninformed on day t, let q_t be the equilibrium fraction that are cherry.

The market price p_t will reflect the uninformed buyers' (rational) beliefs about q_t. That is, p_t is determined by an indifference condition for an uninformed buyer: between holding p_t fruit and purchasing a tree of "quality" q_t. In all the equilibria we examine in sections 3 and 4, we assume that the parameters are such that the informed buyers as a whole cannot afford to purchase all the cherry trees being supplied to the market, so $q_t > 0$.

Two behaviour patterns are obvious. First, a dying trader sells all his trees so as to maximise his fruit consumption before he dies. Second, on the day she is born, an informed trader sells all her endowment of lemon trees: if she uses the proceeds to purchase cherry trees then she is in effect using the market to replace lemon with cherry. (To simplify our diagrams, as an accounting convention we assume that informed new-borns always *first* replace lemon with cherry, even if they have discovered they are going to die later that day and so will immediately resell these cherry trees.)

Finally, we make five assumptions on parameters the reasons for which will not be apparent to the reader until later in the paper:

Assumption 1:

$$0 > a_0 \equiv \theta\mu \left\{ l(1-\gamma)(1-\delta)\delta + l(1-\delta)^2\mu + \Delta\gamma(1-\delta)^2\mu \right\} - \gamma\delta[l\delta - \Delta(1-\delta)\mu] \tag{2.1}$$

Assumption 2:

$$\left[1 - \frac{\delta}{(1-\delta)\mu}\right] - \theta > 0 \tag{2.2}$$

Assumption 3:

$$0 > \tilde{a}_0 \equiv \theta\mu(1-\delta)^2 \left\{(2-\delta)l + (1-\delta)\mu\Delta + \gamma[1+(1-\delta)(1-\mu)]\left[1 - \frac{\delta(1-\mu)}{2-\mu}\right]\Delta\right\}$$

$$- [\gamma\delta - (1-\delta)(1-\gamma)\theta\mu]\left\{\delta(2-\delta)\left[l + \gamma(1-\mu)\left(1 - \frac{\delta(1-\mu)}{2-\mu}\right)\Delta\right] - (1-\delta)^2\mu\Delta\right\} \tag{2.3}$$

Assumption 4:

$$\gamma > \frac{1}{1+(1-\delta)(1-\mu)} \tag{2.4}$$

Assumption 5:

$$0 < \theta\mu(1-\delta)^2 \left\{(2-\mu)(2-\delta)(l+\Delta) - \frac{2(1-\gamma)\Delta[1+(1-\delta)(1-\mu)]^2}{(2-\mu)(1-\delta)}\right\}$$
$$- \left\{\delta^2(2-\delta)(l+\Delta) - \frac{\delta(1-\gamma)\Delta[1+(1-\delta)(1-\mu)]\left[1-(1-\mu)^2(1-\delta)^2\right]}{1-\delta)(1-\mu)(2-\mu)}\right\} \tag{2.5}$$

An analytical proof that parameters consistent with these assumptions exist is left to the appendix.

2.3 Constant Price Equilibrium

In this equilibrium the price and market quality of trees is constant: $p_t = p^*$ and $q_t = q^*$ for all t. Dying traders sell all their trees; new born informed traders switch any lemon trees they carry for cherry (as described in section 2 above) – no other trees are sold. Informed traders exchange all their fruit for cherry trees sold by the dying traders. Uninformed traders purchase the residual stock of trees sold by the dying and the lemons sold by the new-born informed traders. q^* is the proportion of trees bought by the uninformed that are cherry.

2.3.1 Incentive Compatibility and Price

The informed and uninformed equilibrium behaviour follows from the fact that p^* is the value of a market tree to an uninformed trader. The supply of market trees – the residual trees sold after cherry picking by the informed – is positive. Therefore, from the indifference of the uninformed between market trees and p^* fruit, the net purchase of market quality trees by the uninformed is strictly positive i.e., the uninformed do not sell their market quality trees.

Neither do the uninformed sell trees with which they are born as these trees are strictly preferred to p^* fruit. To see this we first derive p^* from its equivalence with an uninformed trader's valuation of a market quality tree:

$$\begin{aligned} p^* &= \mu(l + q^*\Delta) + (1-\mu)p^* \\ &= l + q^*\Delta. \end{aligned} \tag{2.6}$$

Where $(l + q^*\Delta)$ is the expected yield if the tree matures and p^* is the value of the tree if it does not mature.

A proportion γ of trees with which traders are born are cherry – a strictly greater proportion of cherry than for market trees. If a trader holds a tree with which they are born it can be sold for a price p^* if it does not mature. The value of such a tree is then bounded below by

$$\mu(l + \gamma\Delta) + (1-\mu)p^*,$$

which is strictly greater than p^* as defined in equation (2.6). Therefore uninformed traders strictly prefer the trees with which they are born to p^* fruit. The above argument holds *a fortiori* for cherry trees, i.e. informed traders strictly prefer cherry trees to p^* fruit. Therefore informed traders exchange all their fruit for cherry trees and never sell cherry trees if they are not dying that day.

2.3.2 Market Clearing

As the informed traders hold only cherry trees overnight – call this stock N – the uninformed must hold the entire overnight stock of fruit in the economy – call this stock Y. At the equilibrium steady state of Y, Y^*, the total fruit consumed per day equals total fruit produced per day.

The equilibrium steady state of N, N^*, is where the outflow of cherry trees from informed traders' overnight stock is equal to the inflow. Of the $(1-\delta)N^*$ trees held by the non-dying, a fraction μ are lost to maturation. δ of the N^* cherry trees are sold each period by the dying. The maturing cherry trees held by the non-dying informed traders yield $(1-\delta)\mu N^*(l + \Delta)$ fruit – this fruit is then exchanged for cherry trees. The non-dying new-born informed traders replace their lemons with cherry adding $(1-\delta)\theta\mu$ cherry trees to the stock of informed trees each day. The steady state condition for N^* is then:

$$(1-\delta)\mu N^* \quad + \quad \delta N^* \quad = \quad (1-\delta)\mu(l+\Delta)\frac{N^*}{p^*} \quad + \quad (1-\delta)\theta\mu \tag{2.7}$$

Maturing trees held by the non-dying informed	Trees held by the dying informed.	Trees bought from proceeds of maturing trees.	Trees of non-dying new-borns.

We can define all equilibrium stocks and flows in the economy in terms of N^* and Y^*. These stocks and flows are illustrated in figure 1 below:

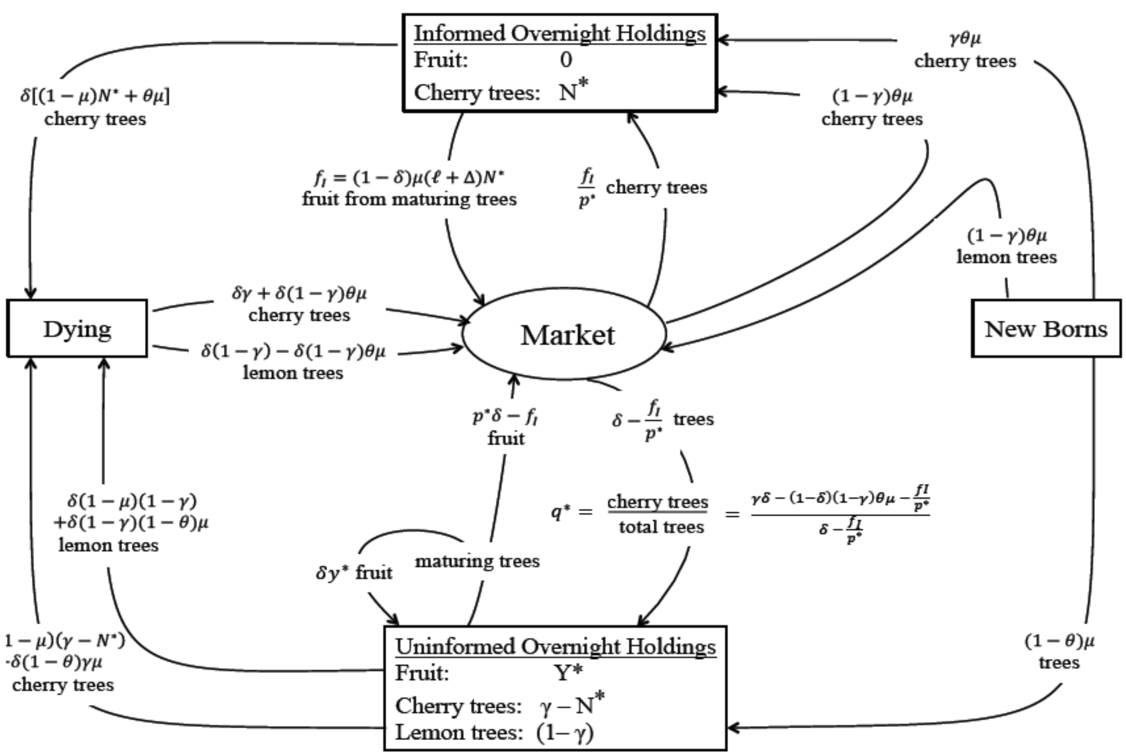


Figure 1: Stocks and Flows in the Constant Price Equilibrium

Markets must clear. In figure 1 we see the total number of trees sold by the dying and the newborn is:

$$\text{Tree Sales} = \underbrace{\delta}_{\substack{\text{From dying}\\ \text{informed and}\\ \text{uninformed.}}} + \underbrace{(1-\gamma)(1-\delta)\theta\mu}_{\substack{\text{Lemon trees from}\\ \text{non-dying}\\ \text{informed new-borns.}}}; \tag{2.8}$$

and the quantity of these tress that are cherry is:

$$\text{Cherry Sales} = \underbrace{\gamma\delta.}_{\substack{\text{Cherry trees from dying}\\ \text{informed and uninformed.}}} \tag{2.9}$$

The informed buy only cherry trees:

$$\text{Informed Cherry Purchases} = \underbrace{(1-\delta)\mu(l+\Delta)\frac{N^*}{p^*}}_{\substack{\text{Cherry trees bought from}\\ \text{proceeds of maturing}\\ \text{cherry trees held by non-}\\ \text{dying informed.}}} + \underbrace{(1-\gamma)(1-\delta)\theta\mu}_{\substack{\text{Cherry trees bought}\\ \text{by non-dying}\\ \text{informed}\\ \text{new-borns.}}} \tag{2.10}$$

The uninformed purchase the residual trees, a fraction of which are cherry:

$$\begin{aligned}\text{Uninformed Cherry Purchases} &= q^* \left\{\underbrace{\delta + (1-\gamma)(1-\delta)\theta\mu}_{\text{Trees sold}} - \underbrace{\left[(1-\delta)\mu(l+\Delta)\frac{N^*}{p^*} + (1-\gamma)(1-\delta)\theta\mu\right]}_{\text{Informed cherry purchases}}\right\} \\ &= q^* \left[\delta - (1-\delta)\mu(l+\Delta)\frac{N^*}{p^*}\right].\end{aligned} \tag{2.11}$$

Hence, the market clearing condition for cherry trees[2] is:

[2]Which also implies market clearing for lemon trees. Total lemon tree sales is: Lemon Sales = Tree Sales − Cherry Sales. Total lemon tree purchases is: Lemon Purchases = $(1-q^*)$[Tree Sales − Informed Cherry Purchases]. Letting Lemon Sales = Lemon Purchases and rearranging yields equation (2.12).

$$\gamma\delta = (1-\delta)\mu(l+\Delta)\frac{N^*}{p^*} + (1-\gamma)(1-\delta)\theta\mu + q^*\left[\delta - (1-\delta)\mu(l+\Delta)\frac{N^*}{p^*}\right]$$

Cherry sales. Informed cherry purchases. Uninformed cherry purchases. (2.12)

Collecting the $\frac{N^*}{p^*}$ terms in the market clearing condition yields:

$$(\gamma - q^*)\delta - (1-\gamma)(1-\delta)\theta\mu = (1-q^*)(1-\delta)\mu(l+\Delta)\frac{N^*}{p^*} \tag{2.13}$$

and doing the same for the steady state stock of informed cherry holdings given by equation (2.7) yields

$$\left\{\frac{l+q^*\Delta}{l+\Delta}[\delta + (1-\delta)\mu] - (1-\delta)\mu\right\}(l+\Delta)\frac{N^*}{p^*} = (1-\delta)\theta\mu \tag{2.14}$$

where $l + q^*\Delta$ comes from equation (2.6) for p^*. Using (2.13) to substitute out $\frac{N^*}{p^*}$ in (2.14) and rearranging gives the quadratic:

$$\begin{aligned}\Phi(q^*) \equiv& (1-\delta)^2\mu^2\theta(1-q^*) \\ &- \left\{\frac{l+q^*\Delta}{l+\Delta}[\delta + (1-\delta)] - (1-\delta)\mu\right\}[\delta(\gamma - q^*) - (1-\gamma)(1-\delta)\theta\mu] \\ =& 0\end{aligned} \tag{2.15}$$

Given the equilibrium behaviour of informed and uninformed traders we require that the solution to this quadratic lies in the interval of $[0, \gamma]$.

The quadratic function $\Phi(.)$ can be written more compactly as:

$$\Phi(q) = a_2 q^2 + a_1 q + a_0; \tag{2.16}$$

where

$$a_2 = \delta\frac{\Delta}{l+\Delta}[\delta + (1-\delta)\mu] > 0. \tag{2.17}$$

The coefficient on the square term, a_2, is positive and a_0 is less than zero by assumption (Assumption 1 section 2). Therefore, if $\Phi(\gamma) > 0$, there exists a unique $q^* \in [0, \gamma]$ such that $\Phi(q^*) = 0$. Define q_N such that:[3]

$$\delta(\gamma - q_N) - (1 - \gamma)(1 - \delta)\theta\mu = 0. \tag{2.18}$$

It is clear that $q_N < \gamma$. We can see that $q_N > 0$ by assumption 2. As $a_0 < 0$ it follows that

$$\begin{aligned} &-l\gamma\delta^2 + \theta\mu l(1 - \gamma)(1 - \delta) < 0 \\ &\Rightarrow \delta\gamma - (1 - \gamma)(1 - \delta)\theta\mu > 0, \end{aligned} \tag{2.19}$$

and therefore:

$$\gamma > q_N > 0. \tag{2.20}$$

Evaluate the sign of $\Phi(.)$ at 0 and q_N:

$$\Phi(0) = a_0 < 0 \tag{2.21}$$

$$\Phi(q_N) = (1 - \delta)^2\mu^2\theta(1 - q_N) > 0. \tag{2.22}$$

Therefore there exists a unique $q^* \in [0, \gamma]$ such that $\Phi(q^*) = 0$ as shown in figure 2 below:

[3] q_N because from (2.13) $N^* = 0$ where $q = q_N$.

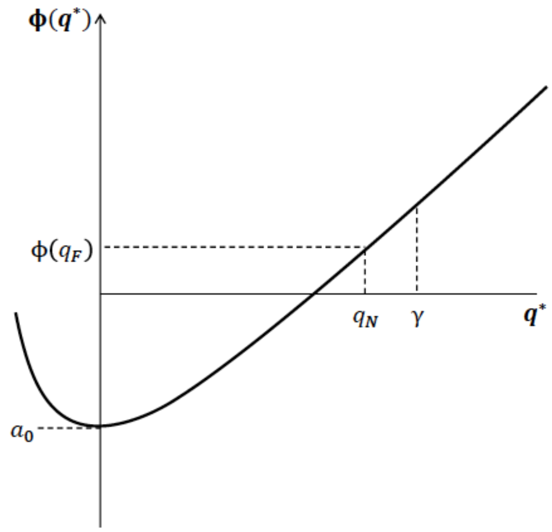


Figure 2

Next, it is shown that $N^* > 0$. Examining the right hand side of the quadratic in (2.15), when $q^* = 0$ the first term is positive and, from equation (2.22), so is the final term in brackets. As $\Phi(0) < 0$ it must be the case that

$$\frac{l + q^*\Delta}{l + \Delta}\left[\delta + (1-\delta)\right] - (1-\delta)\mu > 0. \tag{2.23}$$

Which from equation (2.14) implies that

$$\frac{N^*}{p^*} > 0$$

and therefore $N^* > 0$ for $q^* \in [0, \gamma]$.

The stock of uninformed cherry trees held over night, $(\gamma - N^*)$, must also be greater than zero. In steady state the uninformed lose as many cherry trees as they gain. Hence, $(\gamma - N^*)$ solves

$$(\gamma - N^*)[\delta + (1-\delta)\mu] \quad = \quad \gamma(1-\delta)(1-\theta)\mu \qquad + \quad q^*\left[\delta - (1-\delta)\mu(l+\Delta)\frac{N^*}{p^*}\right]$$

Cherry lost to death and maturation.	Cherry held by undying newborns.	Uninformed cherry purchases.

(2.24)

If the right hand side is positive then $\gamma - N^*$ is positive. The number of cherry trees gained by the uninformed each period is positive if the number of cherry trees purchased by the uninformed is positive. Rearranging the market clearing condition (2.12) gives:

$$(\gamma - q^*)\left[\delta - (1-\delta)\mu(l+\Delta)\frac{N^*}{p^*}\right] = (1-\delta)(1-\gamma)\left[\theta\mu + (l+\Delta)\frac{N^*}{p^*}\right] > 0, \tag{2.25}$$

which implies that uninformed cherry purchases are positive, i.e., their steady state stock of cherry trees is positive.

Finally, we show that the steady state stock of fruit held overnight by the uninformed is positive. The uninformed hold all the fruit in the economy. The steady state stock of fruit in the economy is where total daily fruit consumption equals total daily fruit production.

Dying traders consume their fruit held from the previous day, δY^*; the fruit from selling their trees, δp^*; and the fruit from their maturing trees held from the previous day, $\delta\mu(l+\gamma\Delta)$.[4] Fruit production is due only to maturing trees and is equal to $\mu(l+\gamma\Delta)$. Thus Y^* solves

$$\underbrace{\delta Y^* + \delta p^* + \delta\mu(l+\gamma\Delta)}_{\text{Fruit Consumed by the Dying}} = \underbrace{\mu(l+\gamma\Delta).}_{\text{Fruit from Maturing Trees}} \tag{2.26}$$

Substituting out p^* in (2.26) and rearranging implies that Y^* is positive if:

$$(1-\delta)\mu(l+\gamma\Delta) - \delta(l+q^*\Delta) > 0.$$

$q^* \le \gamma$ hence a sufficient condition for $Y^* > 0$ is:

$$(1-\delta)\mu > \delta,$$

which follows immediately from assumption 2.

2.4 Saw-Tooth Equilibrium

In this equilibrium the market price deterministically oscillates between low on odd days, p_1, and high on even days, p_0. Correspondingly, the quality of trees purchased from the

[4] Fruit is consumed only by the dying. Death is i.i.d. so the average dying trader holds the average of the economy.

market by the uninformed oscillates between low on odd days, q_1, and high on even days, q_0.

On even days all traders sell their trees. All trees sold on even days are purchased by non-dying uninformed traders – no tree is purchased by an informed trader. On even days all trees are sold, the average tree sold is the population average, and informed traders do not cherry pick the market. Hence, the quality of trees purchased by the uninformed on even days is the quality of the population, i.e., $q_0 = \gamma$.[5]

On odd days the dying traders sell all their trees and new born informed traders switch any lemon trees they carry for cherry (as described in section 2) these are the only trees sold on an odd day. Death is i.i.d. so the trees sold by the dying are the population average. The non-dying informed traders exchange all their fruit for cherry trees – cherry picking trees sold by the dying. The uninformed purchase the residual stock of trees sold by the dying and the lemons sold by the new-born informed traders. Hence, the quality of trees purchased by the uninformed on odd days is lower than the population average, i.e., $q_1 < q_0 = \gamma$.

2.4.1 Incentive Compatibility and Prices

Surprisingly, on even days, non-dying informed traders sell trees they know to be cherry to non-dying uninformed traders who only believe the tree is cherry with probability γ. This trade is only possible because of the differing investment horizons of the (non-dying) informed and uninformed traders.

On even days the non-dying informed traders have a shorter investment horizon than the non-dying uninformed: the return to informed traders from purchasing cherry trees is higher on odd days than on even days – so informed traders want to maximise the funds they have available for odd day investments. However, tree holdings suffer a temporary capital loss on odd days when prices drop from p_0 to p_1 – exactly when informed traders need funds for investment. Hence, informed traders are willing to pay a liquidity premium for fruit and take less than the value, $(l+\Delta)$, of a cherry tree's long-run dividend stream.

Non-dying uninformed traders, on the other hand, intend to hold even day trees for the long term – only selling the tree (in effect) when they die. The uninformed traders, if they do not die on the odd day, are unaffected by the temporary capital loss as they (mainly) value the asset for its dividend stream. The longer investment horizon of the uninformed

[5]Note, that the behaviour of the non-dying uninformed is equivalent to holding their population average trees, those they hold from the previous even day or with which they were born, and only selling those trees purchased on odd days.

then allows for gains from trade between the non-dying informed and uninformed traders despite the fact that the informed traders know they are selling their cherry trees for less than the 'true' value.

The volume of trade on odd days is lower than on even days as only the dying and new-born informed with lemons are selling. Trade between the dying and the non-dying occurs due to (the more conventional reason of) different preferences over trees and fruit. Trade between the non-dying informed and uninformed occurs because the informed are selling trees they know to be lemons to traders who believe that the trees are cherry with probability $q_1 > 0$.

Incentive compatibility of the uninformed strategy follows from the non-dying uninformed being indifferent between p_t fruit and a tree of quality q_t at date t. On even days they are indifferent between p_0 fruit and trees of quality γ and so, on even days, must strictly prefer p_0 fruit to trees of quality q_1, so they sell odd day trees on even days. By the same argument, on odd days, they must strictly prefer trees of quality γ to p_1 fruit, so they don't sell population trees on odd days.

Before showing incentive compatibility for the informed it is useful to have expressions for prices on odd and even days. The non-dying uninformed are making speculative purchases with an intent to resell the following day when prices are high. Their hope is that the tree won't mature before they have a chance to sell it and claim the capital gains. Setting p_1 equal to the expected return then yields:

$$p_1 = \mu(l + q_1\Delta) + (1 - \mu)p_0 \tag{2.27}$$

The price on even days is lower than the value, $(l + \gamma\Delta)$, of a population tree's dividend stream because the bearer suffers a capital loss if they die on an odd day before the tree has matured. There are four disjoint outcomes for an uninformed trader holding a population tree:

Day of Event	Event	Expected Payoff	Probability
Odd	Tree Matures	$(l + \gamma\Delta)$	μ
Odd	Tree Doesn't Mature Trader dies	p_1	$(1-\mu)\delta$
Even	Tree Matures	$(l + \gamma\Delta)$	$(1-\mu)(1-\delta)\mu$
Even	Tree Doesn't Mature	p_0	$(1-\mu)^2(1-\delta)$

Setting price equal to expected value yields

$$p_0 = \mu(l + \gamma\Delta) + (1-\mu)\delta p_1 + (1-\mu)(1-\delta)\mu(l+\gamma) + (1-\mu)^2(1-\delta)p_0. \tag{2.28}$$

Substituting equation (2.27) for p_1 in (2.28) and rearranging gives the following expression for price on even days:

$$p_0 = l + \gamma\Delta - (\gamma - q_1)\frac{\delta(1-\mu)}{2-\mu}\Delta. \tag{2.29}$$

Which is the expected value of a population tree's dividend stream less the expected capital loss from having to sell in the event of dying on an odd day. It is worth noting for future reference that both p_0 and p_1 are affine functions of q_1.

It is easy to check that the price does in fact oscillate. From equation (2.27), $p_0 > p_1$ if and only if $p_0 > (l + q_1\Delta)$. Subtracting $l + q_1\Delta$ from (2.29) and rearranging yields

$$p_0 - (l + q_1\Delta) = (\gamma - q_1)\left[1 - \frac{\delta(1-\mu)}{2-\mu}\right]\Delta,$$

which is greater than zero for $q_1 < \gamma$. In any feasible saw-tooth equilibrium the price on even days is higher than the price on odd days.

Returning to incentive compatibility, if an informed trader purchases a cherry tree on an odd day they can always sell on the subsequent day if it does not mature. Hence, the value of a cherry tree to an informed trader on an odd day is bounded below by

$$\mu(l + \Delta) + (1-\mu)p_0 > p_1,$$

i.e. it is always optimal for non-dying informed traders, on odd days, to exchange all their fruit for cherry trees.

From this, on any odd day, informed traders either exchange their portfolio for cherry trees or, if they are dying, exchange their portfolio for fruit, which they then consume. In either event the informed traders utility is linearly increasing in the market value of their portfolio that day. Given, the optimal behaviour on odd days, the non-dying informed traders' decision problem on the even day is to maximise the expected value of their portfolio at price p_1 on the subsequent odd day.

Returns are linear so we can restrict attention to the decision of holding p_0 fruit versus holding a single cherry tree at the end of the even day. The expected market value of a cherry tree held over from the even day is $\mu(l + \Delta) + (1 - \mu)p_1$. Therefore, the informed traders' strategy is incentive compatible if and only if

$$p_0 \geq \mu(l + \Delta) + (1 - \mu)p_1.$$

Using the price equations, (2.27) and (2.29), to substitute out p_0 and p_1 above and rearranging gives the following incentive compatibility condition for informed traders:

$$0 \leq F \equiv -(1 - \gamma) + (\gamma - q_1)(1 - \delta)(1 - \mu). \tag{2.30}$$

We will show later that given the assumptions in section 2 this condition always holds.

2.4.2 Market Clearing

The total equilibrium stocks of fruit held over even nights by informed traders and uninformed traders are W_0 and Y_0 respectively. The total equilibrium stock of cherry trees held over odd nights by informed traders is $\frac{W_1}{p_1}$ and the total equilibrium stock of fruit held over odd nights is Y_1.

Figure 3 shows the stocks and flows of trees and fruit on even days.

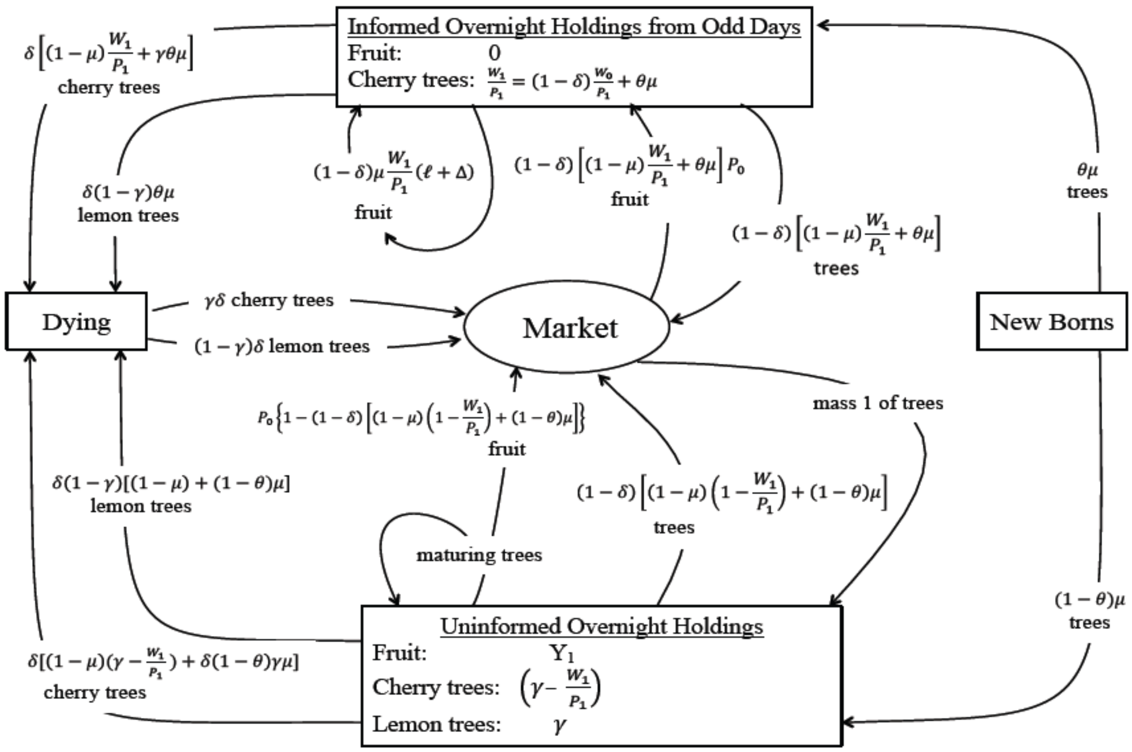


Figure 3: Stocks and Flows on Even Days

Looking at figure 3, the non-dying informed exchange all unmatured cherry trees – both those held over from the odd day and those carried by the new borns – for fruit. The total fruit received is

$$(1-\delta)[(1-\mu)\frac{W_1}{p_1} + \theta\mu]p_0.$$

The non-dying informed also receive

$$(1-\delta)\mu\frac{W_1}{p_1}(l+\Delta)$$

fruit from the cherry trees that matured on the morning of the even day. The informed had no fruit over the odd night and purchased no trees on the even day so the fruit holdings of the informed over even nights is

$$W_0 = (1-\delta)\{\frac{W_1}{p_1}[\mu(l+\Delta) + (1-\mu)p_0] + \theta\mu p_0\}. \tag{2.31}$$

Figure 4 shows the stocks and flows of trees and fruit on odd days.

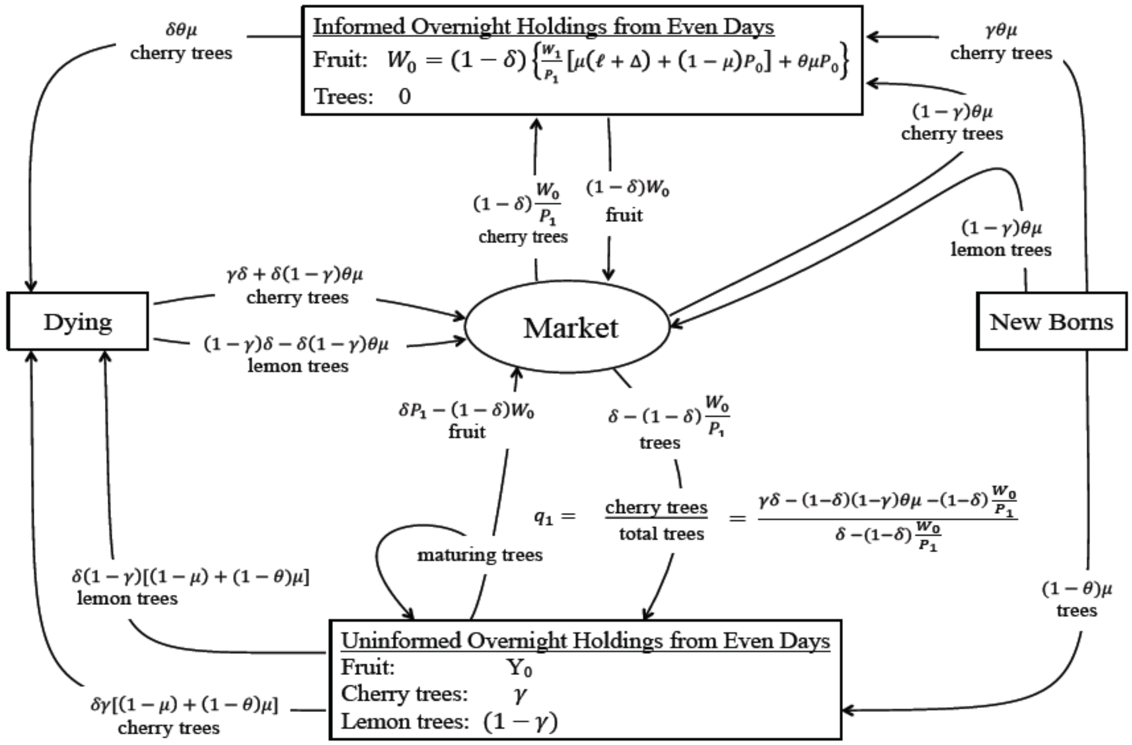


Figure 4: Stocks and Flows on Odd Days

Looking at figure 4, the non-dying informed exchange their fruit stock for

$$(1-\delta)\frac{W_0}{p_1}$$

cherry trees on the odd day; and gain $(1-\delta)\theta\mu$ cherry trees from the non-dying newborn informed, of which a fraction $(1-\gamma)\theta\mu$ were acquired by using the market to replace lemon trees with cherry. The informed cherry tree holdings over odd nights then is

$$\frac{W_1}{p_1} = (1-\delta)(\frac{W_0}{p_1}+\theta\mu). \tag{2.32}$$

The informed hold no fruit over the odd night.

We can now go on to define the market clearing conditions on odd days. In figure 4 we see the total number of trees sold by the dying and the newborn is:

$$\begin{array}{rcll} \text{Tree Sales} & = & \delta & + \quad (1-\gamma)(1-\delta)\theta\mu; \\ & & \begin{array}{l}\text{from dying}\\ \text{informed and}\\ \text{uninformed}\end{array} & \begin{array}{l}\text{lemon trees from}\\ \text{non-dying}\\ \text{informed new-borns}\end{array} \end{array} \tag{2.33}$$

and the quantity of these tress that are cherry is:

$$\begin{array}{rcl} \text{Cherry Sales} & = & \gamma\delta. \\ & & \begin{array}{l}\text{cherry trees from dying}\\ \text{informed and uninformed}\end{array} \end{array} \tag{2.34}$$

The informed buy only cherry trees:

$$\begin{array}{rcll} \text{Informed Cherry Purchases} & = & (1-\delta)\dfrac{W_0}{p_1} & + \quad (1-\gamma)(1-\delta)\theta\mu \\ & & \begin{array}{l}\text{cherry trees bought from}\\ \text{fruit held over even nights}\\ \text{by non-dying}\\ \text{informed}\end{array} & \begin{array}{l}\text{cherry trees bought}\\ \text{by non-dying}\\ \text{informed}\\ \text{new-borns}\end{array} \end{array} \tag{2.35}$$

The uninformed purchase the residual trees, a fraction q_1 of which are cherry:

$$\begin{array}{rl} \begin{array}{l}\text{Uninformed Cherry}\\ \text{Purchases}\end{array} & = q_1 \left\{\delta + (1-\gamma)(1-\delta)\theta\mu - \left[(1-\delta)\dfrac{W_0}{p_1} + (1-\gamma)(1-\delta)\theta\mu\right]\right\} \\ & \qquad\qquad \text{Trees Sold} \qquad - \qquad \text{Informed Cherry Purchases} \\ & = q_1 \left[\delta - (1-\delta)\dfrac{W_0}{p_1}\right]. \end{array} \tag{2.36}$$

Hence, the market clearing condition for cherry trees on odd days[6] is:

[6] Which also implies market clearing for lemon trees. Total lemon tree sales is: Lemon Sales = Tree Sales − Cherry Sales. Total lemon tree purchases is: Lemon Purchases = $(1 - q^*)$[Tree Sales − Informed Cherry Purchases]. Letting Lemon Sales = Lemon Purchases and rearranging yields equation (2.37).

$$\gamma\delta \quad = \quad (1-\delta)\frac{W_0}{p_1} + (1-\gamma)(1-\delta)\theta\mu \quad + \quad q_1\left[\delta - (1-\delta)\frac{W_0}{p_1}\right].$$

Cherry Sales Informed Cherry Purchases Uninformed Cherry Purchases (2.37)

Collecting the $\frac{W_0}{p_1}$ terms in the market clearing condition yields:

$$(\gamma - q_1)\delta - (1-\gamma)(1-\delta)\theta\mu = (1-q_1)(1-\delta)\frac{W_0}{p_1}. \tag{2.38}$$

Solving the steady state equations for W_0 and $\frac{W_1}{p_1}$, (2.31) and (2.32), in terms of $\frac{W_0}{p_1}$ yields:

$$\frac{W_0}{p_1}\left\{p_1 - (1-\delta)^2\left[\mu(l+\Delta) + (1-\mu)p_0\right]\right\}$$

$$= \theta\mu(1-\delta)\left\{(1-\delta)\mu(l+\Delta) + \left[1 + (1-\delta)(1-\mu)\right]p_0\right\} \tag{2.39}$$

Substituting out $\frac{W_0}{p_1}$ using (2.38), and rearranging yields:

$$\begin{aligned}\tilde{\Phi}(q_1) \equiv& (1-\delta)^2(1-q_1)\theta\mu\left\{(1-\delta)\mu(l+\Delta) + [1+(1-\delta)(1-\mu)]p_0\right\} \\ & - \left\{p_1 - (1-\delta)^2\left[\mu(l+\Delta) + (1-\mu)p_0\right]\right\}\left[(\gamma - q_1)\delta - (1-\delta)(1-\gamma)\theta\mu\right] \\ =& 0\end{aligned} \tag{2.40}$$

$\tilde{\Phi}(q_1)$ is quadratic in q_1 as both p_0 and p_1 are affine functions of q_1:

$$\tilde{\Phi}(q_1) = \tilde{a}_2(q_1)^2 + \tilde{a}_1 q_1 + \tilde{a}_0 \tag{2.41}$$

where

$$\begin{aligned}\tilde{a}_2 =& \frac{\delta\mu\Delta}{2-\mu}\left\{(2-\mu) - [1+(1-\delta)(1-\mu)](1-\delta)^2\theta(1-\mu)\right\} \\ &+ \delta^3\Delta(1-\mu)^2\frac{2-\delta}{2-\mu} \\ >& 0 \end{aligned} \tag{2.42}$$

$\tilde{a}_2 > 0$ follows from $[1+(1-\delta)(1-\mu)] < 2-\mu$ and by assumption 4 $\tilde{a}_0 < 0$. As in the constant price equlibrium, $\tilde{\Phi}(0) < 0$ and the co-efficient on the quadratic component is positive so we can show a unique $q_1 \in [0, \gamma]$ solves the quadratic if $\tilde{\Phi}(\gamma) > 0$.

Examining (2.40) it is clear that the argument used in the constant price equilibrium also applies here: $q_1 = q_N \in (0, \gamma)$ and $\tilde{\Phi}(q_N) > 0$ so there exists a unique $q_1 \in (0, \gamma)$ that solves the quadratic.

Examining (2.38) the left hand side is positive where $q_1 \in (0, q_N)$. Therefore where q_1 solves $\tilde{\Phi}(.)$ the fruit holdings of the informed over even nights is positive, $W_0 > 0$. This also implies that the odd night cherry holdings for the informed is positive at the solution as it follows from (2.32) that:

$$(W_0 > 0) \Rightarrow \left(\frac{W_1}{p_1} > 0\right).$$

Over even nights the uninformed hold the entire stock of trees in the economy. Over odd nights the uninformed hold the entire stock of lemon trees and all those cherry trees not held by the informed. The stock of cherry trees held by the uninformed over odd nights is:

$$\underbrace{\gamma - \frac{W_1}{p_1}}_{\text{Uninformed cherry trees held over odd nights}} = \underbrace{(1-\delta)(1-\mu)\gamma}_{\text{Residual stock from even night}} + \underbrace{(1-\delta)\gamma(1-\theta)\mu}_{\text{Endowment of non-dying new-borns}} + \underbrace{q_1\left[\delta - (1-\delta)\frac{W_0}{p_1}\right]}_{\text{Cherry trees purchased by uninformed}} \tag{2.43}$$

The right hand side is positive if the uninformed purchase a positive amount of cherry trees. Rearranging the market clearing condition (2.37):

$$(\gamma - q_1)\left[\delta - (1-\delta)\frac{W_0}{p_1}\right] = (1-\gamma)(1-\delta)\left[\theta\mu + \frac{W_0}{p_1}\right]. \tag{2.44}$$

The right hand side positive so it follows that

$$\delta - (1-\delta)\frac{W_0}{p_1} > 0,$$

and hence cherry purchases of the uninformed on odd days are greater than zero. That is, the uninformed hold a positive stock of cherry trees over odd nights.

The uninformed hold weakly positive stocks of fruit over even and odd nights: $Y_0 > 0$ and $Y_1 > 0$. Over even nights

$$\begin{aligned} Y_0 =& (1-\delta)Y_1 + \\ &+(1-\delta)[\text{ fruit from maturing trees of the uninformed }] \\ &-p_0[\text{Number of trees sold excluding the sales of the non-dying uninformed}] \end{aligned} \tag{2.45}$$

fruit. Before the market on even days the market value of all the trees in the economy not owned by the non-dying uninformed is the expression on the third line.

On the odd days, the non-dying uninformed keep fraction $(1-\delta)$ of the fruit held over the even night. As they hold all trees in the economy over even nights the non-dying uninformed receive a fraction $(1-\delta)$ of all the fruit from maturing trees in the economy. They also purchase all the trees sold by the dying less those bought by the informed. Over odd nights the uninformed fruit holding then is

$$Y_1 = (1-\delta)Y_0 + (1-\delta)\mu(l+\gamma\Delta) - p_1(\delta - (1-\delta)\frac{W_0}{p_1}). \tag{2.46}$$

The uninformed hold over odd nights, as stated earlier, $[(1-\delta)(1-\mu) + (1-\delta)(1-\theta)\mu]$ trees of quality γ and

$$\delta - (1-\delta)\frac{W_0}{p_1}$$

trees of quality q_1. This implies that the uninformed maturing trees yield

$$\mu(1-\delta)[(1-\mu) + (1-\theta)\mu](l+\gamma\Delta) + \mu[\delta - (1-\delta)\frac{W_0}{p_1}](l+q_1\Delta)$$

fruit on even days. Excluding the sales of the non-dying uninformed, the trees sold on even days are those sold by the dying and those sold by the non-dying uninformed, which is a total of

$$\delta + (1-\delta)^2(1-\mu)\frac{W_0}{p_1} + (1-\delta)^2(1-\mu)\theta\mu + (1-\delta)\theta\mu$$

trees.

Using these values and substituting for Y_1 in equation (2.45) using equation (2.46) we get

$$\begin{aligned} Y_0 =& (1-\delta)\{(1-\delta)Y_0 + (1-\delta)\mu(l+\gamma\Delta) - p_1[\delta - (1-\delta)\frac{W_0}{p_1}]\} + \\ &+ (1-\delta)\left\{\mu(1-\delta)[(1-\mu)+(1-\theta)\mu](l+\gamma\Delta) + \mu[\delta-(1-\delta)\frac{W_0}{p_1}](l+q_1\Delta)\right\} \\ &- p_0\left\{\delta + (1-\delta)^2(1-\mu)\frac{W_0}{p_1} + (1-\delta)^2(1-\mu)\theta\mu + (1-\delta)\theta\mu\right\}. \end{aligned} \tag{2.47}$$

Substituting for p_1 above with equation (2.27) and rearranging implies $Y_0 \geq 0$ if and only if

$$\mu(1-\delta)^2(l+\gamma\Delta)(2-\theta\mu) \geq p_0[\delta + (1-\delta)\theta\mu][1 + (1-\delta)(1-\mu)].$$

Given that $p_0 \leq (l+\gamma\Delta)$ a sufficient condition for $Y_0 \geq 0$ is

$$\mu(1-\delta)^2(2-\theta\mu) \geq [\delta + (1-\delta)\theta\mu][1 + (1-\delta)(1-\mu)]$$

or,

$$[1 - \frac{\delta}{\mu(1-\delta)}] - \theta > 0 \tag{2.48}$$

which is true by assumption 2.

Assumption 2 also implies that $Y_1 \geq 0$. The value of $Y_0 \geq 0$ that minimises Y_1 is $Y_0 = 0$. Putting this into equation (2.46)and rearranging yields

$$Y_1 = [(l+\gamma\Delta)(1-\delta)\mu - \delta p_1] + (1-\delta)W_0.$$

Therefore $Y_1 \geq 0$ if $(l+\gamma\Delta)(1-\delta)\mu - \delta p_1$. Given $(l+\gamma\Delta) \geq p_1$ a sufficient condition for $Y_1 \geq 0$ is

$$\mu(1-\delta) > \delta$$

, which is weaker than assumption 2.

It still remains to demonstrate that the equilibrium behaviour for the informed is incentive compatible. In particular that:

$$0 \leq F \equiv -(1-\gamma) + (\gamma - q_1)(1-\delta)(1-\mu), \tag{2.49}$$

where $\tilde{\Phi}(q_1) = 0$ for some $q_1 \in (0, \gamma)$.

F is decreasing in q_1. Define, q_F as that value of q_1 which implies $F = 0$:

$$q_F \equiv q_1 - \frac{(1-\gamma)}{(1-\delta)(1-\mu)}. \tag{2.50}$$

Observe that $q_F > 0$ where

$$\gamma > \frac{1}{1 + (1-\delta)(1-\mu)}$$

which is assumption 4. From rearranging (2.18) and subtracting q_F:

$$q_N - q_F = \frac{(1-\gamma)}{\delta(1-\delta)(1-\mu)} \left[\delta - (1-\delta)^2 \mu(1-\mu)\theta\right].$$

Which is greater than 0 as we know from assumption 1 that $\delta > (1-\delta)^2\mu$.

As $q_N > q_F$ we must demonstrate that $q_1 \in [0, q_F]$ solves $\tilde{\Phi}(q_1) = 0$. This can be done if it is demonstrated that $\tilde{\Phi}(q_F) > 0$, i.e.

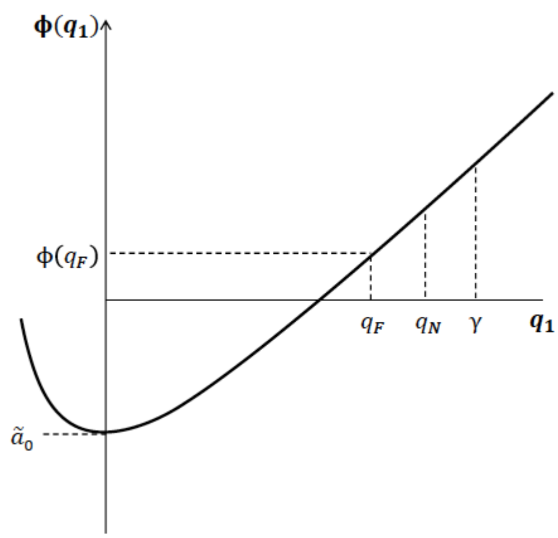


Figure 5

It can be shown that assumption 5 implies $\tilde{\Phi}(q_F) > 0$ and hence there exists a unique incentive compatible $q_1 \in [0, \gamma]$ that solves $\tilde{\Phi}(q_1) = 0$.

2.5 A Numerical Example

Suppose an economy in which 1% of the population is informed; there are 50% more cherry trees than lemon and each cherry tree produces 50% more fruit than a lemon. Finally, one tree matures for every 10 that don't and 1 trader dies for every 20 that don't. Or:

$$\begin{aligned} \mu &= 1/11 \\ \delta &= 1/21 \\ \gamma &= 0.6 \\ \theta &= 0.01 \\ l &= 1 \\ \Delta &= 0.5 \end{aligned}$$

In the constant price equilibrium prices and qualities are:

$$\begin{aligned} q^* &= 0.57 \\ p^* &= 1.28 \end{aligned} \tag{2.51}$$

In the saw-tooth equilibrium prices and qualities are:

$$\begin{aligned} q_0 &= 0.6 \\ q_1 &= 0.12 \\ p_0 &= 1.29 \\ p_1 &= 1.27 \end{aligned}$$

The price on odd days is 2% lower than on even days. From within a completely stationary environment prices cycle deterministically from high to low to high.

Appendix

Proposition 2.1. *There exists a set of parameters that satisfy assumptions 1 through 5.*

To show this set of parameters exist consider the case where $\Delta \to 0$.
Lemma 2.1. *If assumption 3 holds when $\Delta \to 0$ then assumption 1 holds when $\Delta \to 0$.*

Proof:

$$(\Delta \to 0) \Rightarrow a_0 \to l\theta\mu(1-\delta)[(1-\gamma)\delta + (1-\delta)\mu] - l\gamma\delta^2$$

which is less than zero if and only if

$$\gamma\delta^2 > \theta\mu(1-\delta)[(1-\gamma)\delta + (1-\delta)\mu] \tag{2.52}$$

Turn to $\tilde{a}_0$ from Assumption 3:

$$(\Delta \to 0) \Rightarrow \tilde{a}_0 \to (2-\delta)l\{\theta\mu(1-\delta)[1-\delta\gamma] - \gamma\delta^2\}$$

which is less than zero if and only if

$$\gamma\delta^2 > \theta\mu(1-\delta)[1-\delta\gamma]. \tag{2.53}$$

The right hand side of inequality (2.53) is larger than the right hand side of inequality (2.52). Therefore,

$$(\tilde{a}_0 < 0) \Rightarrow (a_0 < 0)$$

□

Lemma 2.2. *There exists a set of parameters consistent with assumption 3 and 5 as $\Delta \to 0$.*

As $\Delta \to 0$ assumption 3 holds if and only if

$$\theta < \frac{\gamma\delta^2}{\mu(1-\delta)(1-\delta\gamma)}. \tag{2.54}$$

As $\Delta \to 0$ assumption 5 tends to

$$0 < (2-\delta)l\left[\theta\mu(1-\delta)^2(2-\mu) - \delta^2\right] \tag{2.55}$$

Hence, assumption 5 holds if and only if

$$\frac{\delta^2}{\mu(1-\delta)^2(2-\mu)} < \theta. \tag{2.56}$$

Therefore, there exist parameters compatible with assumptions 3 and 5 if and only if

$$\frac{\delta^2}{\mu(1-\delta)^2(2-\mu)} < \theta < \frac{\gamma\delta^2}{\mu(1-\delta)(1-\delta\gamma)}. \tag{2.57}$$

(2.57) is possible for any set of parameters satisfying:

$$\gamma > \frac{1}{1+(1-\delta)(1-\mu)} \tag{2.58}$$

□

Note that (2.58) in the proof of lemma 2.2 is assumption 4. Hence assumption 4 is compatible with assumptions 3 and 5. From lemma 2.1 it then follows that there is a set of parameters consistent with assumptions 1, 3,4 and 5.

Finally, we turn to assumption 2:

$$\theta < 1 - \frac{\delta}{(1-\delta)\mu}. \tag{2.59}$$

The lower bound on a value for θ compatible with all five assumptions, as a function of the other parameters when Δ rightarrow 0, is given by assumption 5:

$$\frac{\delta^2}{\mu(1-\delta)^2(2-\mu)} < \theta. \tag{2.60}$$

There exists a θ such that

$$\frac{\delta^2}{\mu(1-\delta)^2(2-\mu)} < \theta < 1 - \frac{\delta}{(1-\delta)\mu} \tag{2.61}$$

when δ is sufficiently small relative to μ. There exists a γ such that assumption 4 is satisfied for any pair of μ and δ. Hence there exists a set of parameters compatible with assumptions 2 and 4. From lemma 2.2 assumption 4 implies there exists parameters that satisfy assumptions 3 and 5 when Δ rightarrow 0. From 2.1 assumption 3 implies assumption 1 when Δ rightarrow 0.

□

Chapter 3

Identifying the Effect of Privilege on Intergenerational Mobility: A Twin Decomposition Method.

James Best

The intergenerational elasticity of income is the standard measurement economists use for intergenerational mobility. It is not clear how we should interpret intergenerational elasticities. Particularly, high intergenerational elasticities could either reflect inequality of opportunity or the importance of genetically heritable characteristics in determining genes. Behavioural geneticists have long been using a twin based variance decomposition method, the ACE model, to estimate the genetic heritability of various characteristics. It is not clear, however, what this approach implies for intergenerational mobility of equality of opportunity.

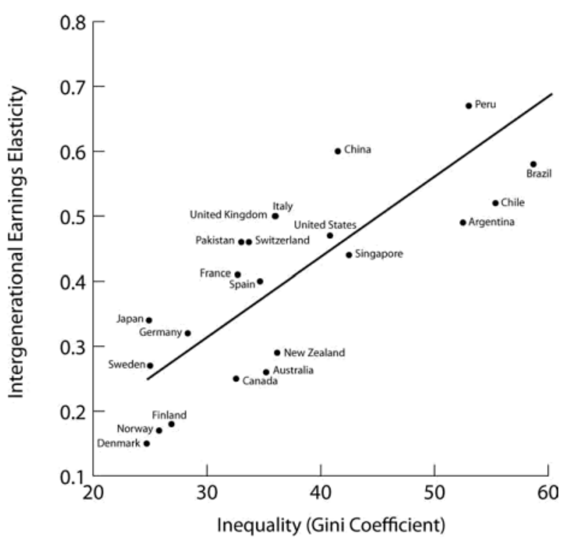
In this chapter I develop a novel method that extends the methodology used in behavioural genetics to identifying how much of the intergenerational elasticity of income is determined by the presence (absence) of environmental privileges associated with being children of high (low) earners. Using this approach we can examine the counterfactuals of giving a poorer child the environment of a richer child; equalising the privileges associated with family income; and equalising the family environmental factors not associated with parental income. Furthermore, this method allows us to identify how good parental income is as a measure of family environment. The model

I develop nests the behavioural genetics model allowing us to relax some of the identifying assumptions used in the standard ACE model.

Finally, I apply this method to data on the income elasticities between American males of different types of relation: fraternal twins, identical twins and father-son relationships. The results of this application suggest that a 1 percent increase in the privilege associated with parental income increases child income by about 1 tenth of a percent. Equalising, to the mean, the environmental privileges across the population results in about a 30 percent drop in the intergenerational elasticity of income and a 5 percent drop in the variance of income across the population. These results must be treated tentatively as the twin data comes from a separate survey to the data on intergenerational elasticities.

3.1 Introduction

High earning parents tend to have high earning children Chetty et al. (2014). The extent of this intergenerational persistence in earnings varies across countries (Jantti et al., 2006; Corak, 2011) and time (Lee and Solon, 2009). There seems, furthermore, to be a relationship between the extent of the intergenerational persistence of earnings in a country and the extent of inequality in a country.



Source: Corak (2011)

This relationship was described by Krueger (2012) as the “Great Gatsby Curve”. This name reflects a concern, voiced by economists and non-economists alike, that income inequality amongst parents is translating into inequality of opportunity for their children and this inequality of opportunity is causing low levels of intergenerational mobility. Put another way, income is determined by privilege and privilege is exacerbated by income inequality. However, an alternative explanation, argued most (in)famously by Herrnstein and Murray (1994), is that the persistence in earnings is due to the genetic transmission of capabilities from parent to child. The policy implications of low intergenerational mobility and high inequality depend greatly on the extent to which these features of the economy are explained by genetics and the extent to which they are explained by variation in the environment provided for children by their families and those around them.

In this paper, I show how to decompose the intergenerational elasticity of income into these two components: the effect of environmental privileges associated with parental income on child outcomes, and the genetic transmission of capabilities between parent and child. I incorporate the standard twin decomposition approach used within behavioural genetics (Neale and Cardon, 1992) into a simple model where there is intergenerational transmission of a characteristic such as income, education or IQ. This method, using data on the income of identical twins, fraternal twins, and their parents can place bounds on the effect of the privileges associated with parental income on inequality and intergenerational mobility. A further contribution of this method is that it relaxes some, but certainly not all, of the assumptions present in the canonical twin decomposition approach.

I apply this method, speculatively, to existent data on income elasticities for American males and their fathers (Lee and Solon, 2009), and for male American identical twins and fraternal twins (Taubman, 1976). The results from this data imply that the privileges associated with paternal income are responsible for about thirty percent of the intergenerational elasticity of income and five percent of the variation in income. The fact that the income elasticities between fathers and sons come from a different source, and cohort, to those between different types of twins is a potential source of bias. Consequently, too much weight should not be put on the empirical results reported in this paper. These results are useful, however, for demonstrating this paper's methodological contribution and its implications.

The standard twin decomposition approach, or ACE model, uses data on the correlation between identical twins (raised in the same family) and fraternal twins (raised in the same family). This approach takes advantage of the fact that identical twins have all their genes in common while fraternal (non-identical) twins do not. Particularly, when there is random mating, Mendel's laws of genetics imply the correlation between the genetic polymorphisms of fraternal twins is half. Using this fact, and other assumptions to be discussed later, it can be estimated how much of the variation in some characteristic is explained by 'additive genetics', by 'common environment' and by 'unique environment'. Where common environment refers to the environmental elements which twins share due to being born and raised in the same family, for this reason it is often referred to as 'family environment'. Unique environment refers to that part of a twin's environment which is uncorrelated with their genes.

The headline result of the ACE model is that the proportion of variance explained by genes, typically denoted as a^2 for 'additive genetics', is twice the difference in the correlation of identical twins and fraternal twins. Suppose the outcome of interest were IQ then

$$a^2 = 2\left[\mathrm{Corr}(IQ_1, IQ_2|\text{identical twins}) - \mathrm{Corr}(IQ_1, IQ_2|\text{fraternal twins})\right].$$

The proportion of variance explained by family environment, typically denoted as c^2 for 'common environment', is the difference between twice the correlation of fraternal twins and the correlation of identical twins, so for IQ this would be

$$c^2 = 2\,\mathrm{Corr}(IQ_1, IQ_2|\text{fraternal twins}) - \mathrm{Corr}(IQ_1, IQ_2|\text{identical twins}).$$

Finally, the proportion explained by the unique environment of a twin, e^2, is that part of variance not explained by either genes or family.

The ACE model, and variants thereof, have been used by behavioural geneticists to examine IQ, education, personality, willpower, and many more characteristics since at least Merriman (1924). Turkheimer (2000) states that the first two laws of behavioural genetics are 1) all human behavioural characteristics are heritable; and 2) the effect of being raised in the same family is smaller than the effect of genes. A review by Boomsma et al. (2002) reports that the effect of being raised in the same family is zero for many behavioural traits, including the IQ of adults. There have also been some studies by economists which examine income directly. The main one[1] examining the income of American twins is Taubman (1976). He finds that genes explain 48% and family environment explains 6% of the variance in income.

The above findings of behavioural geneticists and economists using the classic twin study seem to lend some *prima facie* support to the argument that the persistence of earnings and inequality in earnings is largely due to genes. However, these results do not give the effect on children of the environmental advantages associated with parental income. Consequently, the implications of these results for intergenerational mobility are unclear. Nor do these results tell us the effect on inequality of removing the environmental advantages associated with higher parental incomes. Goldberger (1979) and Manski (2011) are

[1] There is later work using the same twin data set but the methodology and results are very similar.

particularly scathing about the policy relevance of these results.

The reason these models are unable to provide answers to these questions is because they are silent on what constitutes the 'common environment' of twins. By incorporating these models into a model of intergenerational mobility we can examine the effects of that part of family environment which is explained by parental income and its correlates. This also allows us to evaluate the importance in determining children's outcomes of those features of the family environment which are unassociated with (are orthogonal to) parental income[2]. Most importantly, however, we can run the counterfactual experiment of equalising the environmental advantages, privileges, associated with parental income.

The standard ACE model also requires a series of covariance restrictions in order to identify the relevant parameters. One of these restrictions is that the family environment is uncorrelated with genes. This restriction is very strong in the case of many characteristics. Take income as the point in case, where the estimate of heritability of income is substantial. For genes to be uncorrelated with family environment in this case one of the following must be true: a) parental income is not a relevant part of shared environment; b) child income is heritable parental income is not; or c) those genes relevant to parental income are no longer relevant for child income.

"The Bell Curve" (Herrnstein and Murray, 1994) while the most well known argument for the heritability hypotheses of intergenerational mobility has many flaws. Bowles et al. (2001) and Bowles and Gintis (2002) are more thorough in examining how intergenerational mobility is explained by IQ and other heritable measures. They find that the heritability of IQ explains little of the intergenerational persistence in earnings. Their result is driven chiefly by the fact that IQ does not explain a large proportion of income variation. Further work by Blanden et al. (2007) examine the effect of cognitive and non-cognitive skills. They find that they account for about half the intergenerational elasticity. They, contra Bowles et al. (2001), do not do not access the potential role for genes as a mechanism by which cognitive and non-cognitive skills are transmitted from parents to children.

Björklund and Jäntti (2009) is the closest research to my own. They use excellent Swedish data in which there are many different types of siblings to relax and directly estimate some of the correlations in environment and genes between different types of relations. This is similar to the method used for relaxing the restriction on the correlation between genes and family environment in this paper. However, they do not apply their methodology to examining intergenerational mobility or estimating the environmental impact of privileges associated with parental income.

[2]Note that this gives an estimate of the extent of measurement error of parental income as a measure of family investment.

There is also a literature on estimating the returns to education using identical twins such as that of Krueger and Ashenfelter (1994) and Isacsson (1999). The identical twin approach rests on an assumption that identical twins have identical ability and therefore the effects of any variation in schooling can be taken as the returns to schooling. For a critique of this approach see Bound and Solon (1999). This approach is completely different to the variance decomposition approach applied in the ACE model. One of the many large differences is that they answer different questions. The ACE model tells us about the relative importance of variation across family environment in explaining the population variation in some particular outcome. The identical twin approach used for estimating the returns to education examines the effect of some particular environmental factor that varies within a family.

There are reasons to be sceptical of the assumptions underlying the ACE model. There are issues regarding assortative mating, gene-environment interactions, twin interactions, non-linearities in the effects of genes and correlation between family environment and genes that will be discussed in more detail later in the paper. The method I develop allows us to deal with one of these issues, the assumption of family environment being uncorrelated with genes. Also it allows us to interpret the twin study results in a way that is meaningful for the debate over inequality and intergenerational mobility.

In section II, I introduce the ACE model and its identifying assumptions. In section III, I introduce my model in which the classic twin approach is embedded in a simple model of intergenerational mobility and show how we can use it to identify the effect of privilege on child outcomes. In section IV, I describe several counterfactual experiments and how my model can be used to estimate the effects. In section V, I apply my method to existent results on twin elasticities and intergenerational elasticities. In this section I also give the results of the counterfactual experiments described in section IV. In section VI, I discuss some of the limitations of the approach and the data used in section V.

3.2 The Behavioural Genetics Benchmark

3.2.1 The ACE Model

The ACE model is the benchmark model used within behavioural genetics: ‘A’ stands for ‘additive genetics’; ‘C’ for common environment; and ‘E’ for unique environment. For a twin, in some family i, A_i refers to their genetic endowment. Common environment, C_i, refers to those environmental factors that they share with their twin. Unique environment, E_i, refers to those environmental factors that are unique to a twin, i.e. not shared with their twin, and are also independent of their genes.

It is assumed that the effect of genes on some outcome are additively separably from the effects of environment. Moreover, the effect of any particular genetic polymorphism is independent of the presence of other genes. In an additive genetic model we can think of the effect of genes on some outcome as simply the sum of the effects of different genetic polymorphisms

$$\sum_{l=1}^{L} b_l g_{i,l}.$$

Where $g_{i,l}$ is a dummy variable for the presence of genetic polymorphism l and b_l is the effect of polymorphism l on the outcome ω_l, i.e.,

$$\frac{d\omega_i}{dg_{i,l}} = b_l.$$

The ACE model does not make use of directly measured genetic polymorphisms[3], i.e. measurements of g_l or $\sum_{l=1}^{L} b_l g_{i,l}$ are not included in the model. Instead the genetic endowment is treated as a latent, that is unobserved, factor. Our knowledge of Mendelian genetics can, conditional on the modelling assumptions, be then used to infer how important genes are in explaining the variation of some outcome in the population, and by extension, to explain the importance of environment.

Let $\omega_i^{j,r}$ be the outcome of interest in family i for twin j in a twin pair with relatedness $r \in \{mz, dz\}$. Relatedness, $r \in \{mz, dz\}$, denotes whether the twins are monozygotic, mz, or dizygotic, dz. Recall that monozygotic twins are identical twins and dizygotic twins are fraternal, i.e. non-identical, twins. $\omega_i^{j,r}$ is determined by $A_i^{j,r}$, $C_i^{j,r}$ and $E_i^{j,r}$.

To simplify notation the superscripts are dropped when we are only considering an individual and not their twin. ω_i, A_i, C_i and E_i are all normalised so that they are distributed in the population with mean 0 and variance 1. The outcome of an individual is then

$$\omega_i = aA_i + cC_i + eE_i \tag{3.1}$$

[3]Recently, technology has been developed that can directly measure genetic polymorphisms relatively cheaply. Unfortunately the number of polymorphisms for any individual dwarf the number of people within the samples and, at this time, attempts to directly estimate the effects of genes on complex traits, such as IQ or income, have been largely unsuccessful. There has been some research finding that some particular genetic polymorphism has a tiny effect on complex traits such as IQ, however, when these tests have been replicated the effects have disappeared.

3.2.2 Identification

A_i, C_i and E_i, being latent factors, are unobserved and hence can be normalised so that their respective population distributions have mean zero and variance one. To identify a, c and e behavioural geneticists impose a set of restrictions on the variance-covariance structure between the factors of individuals; between the factors of identical twins; and between the factors of fraternal twins.

Before proceeding further it is useful to establish some notation. The correlation in outcome, ω_i, for identical twins is denoted as

$$\mu = \text{Corr}(\omega_i^{1,mz}, \omega_i^{2,mz});$$

and for fraternal twins as

$$\delta = \text{Corr}(\omega_i^{1,dz}, \omega_i^{2,dz}).$$

μ is for 'monozygotic' and δ is for 'dizygotic'. The correlation in some factor X for identical twins is denoted as

$$\mu_X = \text{Corr}(X_i^{1,mz}, X_i^{2,mz});$$

and for fraternal twins as

$$\delta_X = \text{Corr}(X_i^{1,dz}, X_i^{2,dz}).$$

When the correlation is between a factor X and a factor Y then the correlation for identical twins is denoted by

$$\mu_{XY} = \text{Corr}(X_i^{1,mz}, Y_i^{2,mz});$$

for fraternal twins by

$$\delta_{XY} = \text{Corr}(X_i^{1,dz}, Y_i^{2,dz});$$

and for individuals by

$$\iota_{XY} = \text{Corr}(X_i, Y_i).$$

ι is for 'individual'. Finally, let

$$\boldsymbol{\alpha} = \begin{pmatrix} a \\ c \\ e \end{pmatrix} \quad \text{and} \quad \mathbf{y}_i^{j,r} = \begin{pmatrix} A_i^{j,r} \\ C_i^{j,r} \\ E_i^{j,r} \end{pmatrix}$$

We can now return to the specifics of the restrictions imposed on the relevant variance covariance matrices. The identifying restrictions on the factor variance-covariance structure for individuals imply an identity matrix:

$$E[\mathbf{y}_i\mathbf{y}_i'] = \begin{pmatrix} 1 & & \\ \iota_{AC} & 1 & \\ \iota_{AE} & \iota_{CE} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

E is defined as that part of a twins environment which they do not share with their twin and that is also uncorrelated with their genes. Hence, $\iota_{AE} = \iota_{CE} = 0$ by definition. However, $\iota_{AC} = 0$ only by assumption. That is it is assumed that family environment is uncorrelated with genes. As discussed in the introduction this is a strong and implausible assumption that my method allows us to relax.

The identifying restrictions on the factor variance-covariance structure between identical twins are

$$E[\mathbf{y}_i^{1,mz}(\mathbf{y}_i^{2,mz})'] = \begin{pmatrix} \mu_A & & \\ \mu_{AC} & \mu_C & \\ \mu_{AE} & \mu_{CE} & \mu_E \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

and between fraternal twins they are

$$E[\mathbf{y}_i^{1,dz}(\mathbf{y}_i^{2,dz})'] = \begin{pmatrix} \delta_A & & \\ \delta_{AC} & \delta_C & \\ \delta_{AE} & \delta_{CE} & \delta_E \end{pmatrix} = \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The correlation of 1 for the genes of identical twins follow because their genes are identical. The restriction of the correlation of a half for the genes of the fraternal twins, $\delta_A = 0.5$, relies on Mendel's first and second law and an assumption of random mating.

Many of the correlations involving the unique environment factor follow by definition. First, if it were not the case that $\mu_E = \delta_E = 0$ then $E_i^{j,r}$ would not be unique to twin j. Second, the common environment of twins is identical i.e.$C_i^{1,r} = C_i^{2,r}$. It follows from this that if one twin's unique environment, $E_i^{j,r}$, is uncorrelated with their common environment, $C_i^{j,r}$, it must also be uncorrelated with their twins common environment, $C_i^{j,r}$ – that is $\mu_{EC} = \delta_{EC} = 0$.

It may also seem, at first glance, that the definition of common environment, $C_i^{1,r} = C_i^{2,r}$, also implies that $\mu_C = \delta_C = 1$. However, the normalisation of both $C_i^{j,mz}$ and $C_i^{j,dz}$ to have $\text{Var}(C_i^{j,mz}) = \text{Var}(C_i^{j,dz}) = 1$ is an assumption that the common environment component is as important for identical twins as it is for fraternal twins. This is often, and misleadingly, referred to as the equal environment assumption.

While there are valid concerns over the equal environment assumption let's examine a particularly common criticism which is misconceived. A persons genes evoke different environmental responses and can also cause them to actively seek out different environments. For example, people may be kinder to more beautiful people or people who are naturally bright may seek out more stimulating environments. Identical twins share more genes in common than fraternal twins and so the environments that they evoke and seek out will be more similar. Consequently, it is reasonable to say that identical twins share more environment in common than do identical twins. Thus, so the argument goes, the equal environment assumption is deeply flawed.

This argument is flawed because it misinterprets how genes, A_i, affect the outcome of interest. The estimate of the effect of genes on the outcome includes two kinds of effects of a child's genes on their own environment: 1) that their genes might evoke particular environments e.g. beautiful children are treated more kindly; and 2) that their genes might cause them to seek out particular environments e.g. naturally bright children seek out libraries. In the case of income, part of the effect of genes on income is that those genes help lead to environments that are more conducive to earning. Hence, correlation in environment caused by the mechanisms of evocation and active seeking discussed above are not breaches of the so-called equal environment assumption.

Next, consider the assumption that $\mu_{AE} = \delta_{AE} = 0$. It is clear that $\mu_{AE} = 0$ for the same reason that $\iota_{AE} = 0$: identical twins have identical genes and so if one twin's genes are uncorrelated with their unique environment their genes must also be uncorrelated with their twins unique environment. However, those genes a fraternal twin does not share in common with their twin could be correlated with their twins unique environment. One potential cause of this correlation is the fact that part of a twins' environment is the genes of their twin. The effect of this should be to make identical twins share more environment than fraternal twins and the size of this effect depends greatly on how important a person's twin is for their own development.

Finally, it follows from the definition of common environment that $\iota_{AC} = \mu_{AC} = \delta_{AC}$. Hence, the assumption of zero correlation between one twin's common environment and the other twin's genes stands on the same (shaky) ground as the assumption of $\iota_{AC} = 0$. My method, as will be seen, allows us to relax these assumptions.

Given these restrictions, the variance in outcomes is

$$\text{Var}[\omega_i] = 1 = \boldsymbol{\alpha}' E[\mathbf{y}_i \mathbf{y}_i'] \boldsymbol{\alpha} = a^2 + c^2 + e^2;$$

the correlation between identical twins is

$$\mu = \boldsymbol{\alpha}' E[\mathbf{y}_i^{1,mz}(\mathbf{y}_i^{2,mz})']\boldsymbol{\alpha} = a^2 + c^2;$$

and the correlation between fraternal twins is

$$\delta = \boldsymbol{\alpha}' E[\mathbf{y}_i^{1,dz}(\mathbf{y}_i^{2,dz})']\boldsymbol{\alpha} = 0.5a^2 + c^2.$$

From these three moment conditions we can identify all three parameters of interest:

$$a^2 = 2(\mu - \delta), \tag{3.2}$$
$$c^2 = 2\delta - \mu, \tag{3.3}$$
$$e^2 = 1 - \mu. \tag{3.4}$$

The interpretation of a^2, c^2 and e^2 is the amount of variance in the outcome ω_i that is explained by additive genetics, shared environment and unique environment respectively.

Given the identifying assumptions this model tells us what the effect on inequality is of equalising the common environment of all twins. Particularly, it will reduce the variance in income, or some other outcome of interest, by c^2. We can also ask about the effect of moving a child up one standard deviation in the distribution of common environment, in expectation it would improve their outcome by c. However, this model is silent about what constitutes common environment or what it would mean to move a child one standard deviation in the distribution of common environment. Consequently, the results from the ACE model does not identify the environmental effect of parental income and its associated privileges on the income of children. With this in mind we can turn to the intergenerational twin model which does allow us to identify this effect.

3.3 An Intergenerational Twin Model

I incorporate the standard ACE model into an intergenerational framework. This allows us to examine the implications of the standard ACE model for the mechanisms which cause intergenerational mobility, or lack there of. This method generalizes to the intergenerational persistence of other characteristics such as education, IQ, unemployment, criminality and so on. In the process of embedding the ACE model within an intergenerational framework we are also able to relax some of the identifying assumptions in the standard behavioural genetics model.

This model continues to use several assumptions from the behavioural genetics model. Particularly, the additive genetics assumption; the random mating assumption; the equal environment assumption; and the assumption that the genes of one twin are uncorrelated with the unique environment of the other. It is no longer assumed, however, that shared environment is uncorrelated with genes:

$$\iota_{AC} \neq 0$$

is possible.

3.3.1 Model

Factors are given a generational subscript, n, where generation n are the children of generation $n-1$. For example, $\omega_{i,n}^{j,r}$ is the outcome of interest for twin j in family i of generation n and $\omega_{i,n-1}$ is the outcome of interest for that twin's father or mother. Note that the father is not necessarily (probably isn't) a twin, hence the absence of superscripts.

As in the ACE model $\omega_{i,n}$ is taken to be determined by a latent, i.e. unobserved, genetic factor, $A_{i,n}$, and a latent unique environment factor, $E_{i,n}$. The common environment factor, however, is decomposed into two parts: a part explained by parental income and a part which is orthogonal to parental income. That part that is explained by parental income is referred to from here on out as 'privilege'. This factor is denoted as $P_{i,n}$. That part which is orthogonal to parental income is (somewhat clumsily) referred to as 'residual family environment'. This factor is denoted as $F_{i,n}$. As $F_{i,n}$ is independent of parental income it is also independent of privilege:

$$P_{i,n} \perp\!\!\!\perp F_{i,n}.$$

Both privilege, $P_{i,n}$, and the residual family environment, $F_{i,n}$, are normalised to have mean 0 and variance 1. The effect of common environment in the ACE model is then the combined effect of privilege and residual family environment in the intergenerational ACE model:

$$cC_{i,n} = pP_{i,n} + fF_{i,n}. \tag{3.5}$$

p is the environmental impact of privilege, i.e. the environmental impact of parental income and its correlates. Where as f is the environmental impact of the relevant components of environment shared by both twins that are uncorrelated with parental income.

Note that p captures part of parental investment, in a reduced form way. Income is an imperfect measure of parental investment, see Heckman and Mosso (2014) for a richer structural model of investment, but by identifying p and f we can better understand the relationship between parental income and parental investment. It follows from (3.5) and the independence of $P_{i,n}$ and $F_{i,n}$ that:

$$\text{Var}(cC_{i,n}) = c^2 = p^2 + f^2.$$

Hence, parental income explains

$$\frac{p^2}{p^2 + f^2} \tag{3.6}$$

of the common environment. Or put slightly differently, the correlation between parental income and common environment is:

$$\text{Corr}(\omega_{i,n-1}, C_{i,n}) = \frac{p}{\sqrt{p^2 + f^2}}. \tag{3.7}$$

We can embed the ACE model within an intergenerational framework by using (3.5) to substitute out for $cC_{i,n}$ in the standard ACE model:[4]

$$\omega_{i,n} = aA_{i,n} + pP_{i,n} + fF_{i,n} + eE_{i,n} \tag{3.8}$$

The intergenerational transmission of characteristics is characterised as a stationary process. Let the vector of factors determining the characteristic of interest be denoted as:

$$\mathbf{z}_{i,n} = \begin{pmatrix} A_{i,n} \\ P_{i,n} \\ F_{i,n} \\ E_{i,n} \end{pmatrix}$$

The evolution of $\mathbf{z}_{i,n}$ overtime is then a stationary vector auto-regressive factor model:

[4] It is important to realise that as $P_{i,n}$ and $\omega_{i,n-1}$ are perfectly correlated that $P_{i,n}$ in (3.8) could be replaced by $\omega_{i,n-1}$. Privilege, $P_{i,n}$, is only used instead of parental income, $\omega_{i,n-1}$, to allow clarity of exposition. Particularly, it is useful to keep parental income distinct from the environmental impact of parental income and its correlates on children.

$$\mathbf{z}_{i,n} = \mathbf{B}\mathbf{z}_{i,n-1} + \mathbf{v}_{i,n} \quad \text{where} \quad \mathbf{B} = \begin{pmatrix} \mathbf{b}_1' \\ \boldsymbol{\beta}' \\ \mathbf{b}_3' \\ \mathbf{0}' \end{pmatrix} = \begin{pmatrix} \phi_{AA} & \phi_{A\omega} & \phi_{AF} & \phi_{AE} \\ a & p & f & e \\ \phi_{FA} & \phi_{F\omega} & \phi_{FF} & \phi_{FE} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\text{and} \qquad \mathbf{v}_{i,n} = \begin{pmatrix} \epsilon_{i,n} \\ 0 \\ \xi_{i,n} \\ \eta_{i,n} \end{pmatrix} \perp\!\!\!\perp \mathbf{z}_{i,n-1}$$

The outcome of interest as a function of these four factors is:

$$\omega_{i,n} = \boldsymbol{\beta}'\mathbf{z}_{i,n} \tag{3.9}$$

Hence, the outcome of interest is

$$\omega_{i,n} = \boldsymbol{\beta}'\mathbf{z}_{i,n} = \boldsymbol{\beta}'\mathbf{z}_{i,n-1} + \boldsymbol{\beta}'\mathbf{v}_{i,n}. \tag{3.10}$$

3.3.2 Identification

Bounds can be put on all parameters in $\boldsymbol{\beta}$ using 4 population moments, a stationarity assumption, and a strict subset of the identifying restrictions used within the classic twin decomposition model used in behavioural genetics. The moments are the covariance between the outcomes of identical twins, $\text{Cov}(\omega_{i,n}^{1,mz}, \omega_{i,n}^{2,mz})$; the covariance between the outcomes of fraternal twins, $\text{Cov}(\omega_{i,n}^{1,dz}, \omega_{i,n}^{2,dz})$; the covariance between parent outcome and child outcome, $\text{Cov}(\omega_{i,n}, \omega_{i,n-1})$; and the variance of the outcome of interest, $\text{Var}(\omega_{i,n}) = 1$.

The notational conventions used in the ACE model are continued. There is some additional notation. The intergenerational correlation between the outcome of the parent, $\omega_{i,n-1}$, and the outcome of the child, $\omega_{i,n}$ is denoted as

$$\gamma = \text{Corr}(\omega_{i,n-1}, \omega_{i,n}).$$

γ is for 'generation'. The intergenerational correlation of some variable X other than the outcome of interest is denoted as

$$\gamma_X = \text{Corr}(X_{i,n-1}, X_{i,n}).$$

Finally, the correlation between the factor X of the parent and the factor Y of the child is

$$\gamma_{XY} = \text{Corr}(X_{i,n-1}, Y_{i,n}).$$

Note that the order of subscripts for intergenerational correlations, contra individual and twin correlations, has significance. Particularly, the first subscript relates to the parent's factor and the second subscript to the child's factor.

The primary goal of this exercise is to decompose the intergenerational elasticity of income into a component explained be genetics and a component explained by the environmental impact of environmental correlates of parental income. It is easy to show that this decomposes as follows:

$$\gamma = \underbrace{a\gamma_{\omega A}}_{\text{genetic component}} + \underbrace{p}_{\text{environmental component}} \tag{3.11}$$

Recalling that the income of the child is a function of child characteristics

$$\omega_{i,n} = \boldsymbol{\beta}' \mathbf{z}_{i,n}.$$

the correlation of child and parent income is:

$$\begin{aligned}\gamma =& E[\omega_{i,n-1}\boldsymbol{\beta}' \mathbf{z}_{i,n-1}] \\ =& a\gamma_{\omega A} + p\gamma_{\omega P} + f\gamma_{\omega F} + e\gamma_{\omega E} \\ =& a\gamma_{\omega A} + p\end{aligned}$$

From the definition of $F_{i,n}$ and $E_{i,n}$ they are uncorrelated with parental income. While from the definition of privilege $P_{i,n}$ is perfectly correlated with parental income and so $\gamma_{\omega P} = 1$.

We continue with the standard behavioural genetics assumption that the unique environment of one twin is uncorrelated with the genes of the other twin. Also recall that, by definition, the unique environment of one twin is uncorrelated with both the unique environment and shared environment of the other twin. These assumptions and definitions imply that moment conditions for the correlation between the outcome of the identical twins, μ, and the outcome of the fraternal twins, δ are as follows:

$$\begin{aligned}\mu =&E[\boldsymbol{\beta}'\mathbf{z}_{i,n}^{1,mz}(\mathbf{z}_{i,n}^{2,mz})'\boldsymbol{\beta}]\\ =&a^2\mu_A + p^2\mu_P + f^2\mu_F + a\left\{pE\left[P_{i,n}(A_{i,n}^{1,mz} + A_{i,n}^{2,mz})\right] + fE\left[F_{i,n-1}(A_{i,n}^{1,mz} + A_{i,n}^{2,mz})\right]\right\}\\ =&a^2\mu_A + p^2\mu_P + f^2\mu_F + 2a(p\iota_{AP} + f\iota_{AF}),\end{aligned}$$

and

$$\begin{aligned}\delta =&E[\boldsymbol{\beta}'\mathbf{z}_{i,n}^{1,dz}(\mathbf{z}_{i,n}^{2,dz})'\boldsymbol{\beta}]\\ =&a^2\delta_A + p^2\delta_P + f^2\delta_F + a\left\{pE\left[P_{i,n}(A_{i,n}^{1,dz} + A_{i,n}^{2,dz})\right] + fE\left[F_{i,n-1}(A_{i,n}^{1,mz} + A_{i,n}^{2,mz})\right]\right\}\\ =&a^2\delta_A + p^2\delta_P + f^2\delta_F + 2a(p\iota_{AP} + f\iota_{AF}).\end{aligned}$$

Finally, the moment condition for the variance of ω is:

$$\begin{aligned}\text{Var}(\omega_{i,n}) = 1 =&E[\boldsymbol{\beta}'\mathbf{z}_{i,n}\mathbf{z}_{i,n}'\boldsymbol{\beta}]\\ =&a^2 + p^2 + f^2 + e^2 + 2a\left\{pE\left[A_{i,n}P_{i,n}\right] + fE\left[A_{i,n}F_{i,n}\right]\right\}\\ =&a^2 + p^2 + f^2 + e^2 + 2a(p\iota_{AP} + f\iota_{AF}).\end{aligned}$$

To identify the four parameters of interest, values for $\gamma_{\omega A}$, ι_{AP} and ι_{AF} are required. By definition

$$\gamma_{\omega A} = \iota_{AP}.$$

ι_{AP} can then be derived as a function of the four parameters of interest, $\mathbf{b}_1$ and ι_{AF}:

$$\begin{aligned}\iota_{AP} =&E[P_{i,n}A_{i,n}] = E[\omega_{i,n-1}A_{i,n}]\\ =&E[(\boldsymbol{\beta}'\mathbf{z}_{i,n-1})(\mathbf{b}_1'\mathbf{z}_{i,n-1} + \epsilon_{i,n})]\\ =&\boldsymbol{\beta}'E[\mathbf{z}_{i,n-1}\mathbf{z}_{i,n-1}']\mathbf{b}_1 + E[(\boldsymbol{\beta}'\mathbf{z}_{i,n-1})\epsilon_{i,n}]\\ =&\boldsymbol{\beta}'E[\mathbf{z}_{i,n-1}\mathbf{z}_{i,n-1}']\mathbf{b}_1\end{aligned}$$

The variance-covariance structure of parental characteristics which is:

$$E[\mathbf{z}_{i,n-1}(\mathbf{z}_{i,n-1})'] = \begin{pmatrix} 1 & & & \\ \iota_{AP} & 1 & & \\ \iota_{AF} & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

To proceed further we must reintroduce the behavioural genetics assumptions. For the rest of this paper we will take Mendel's first two laws and the assumption of random mating in the behavioural genetics models as given. This implies that the first row, $\mathbf{b}_1'$, of the transmission matrix, $\mathbf{B}$, is:

$$\mathbf{b}_1 = \begin{pmatrix} 0.5 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Hence, under these assumptions ι_{AP} is given as

$$\iota_{AP} = 0.5[a + p\iota_{AP} + f\iota_{AF}] \tag{3.12}$$

Evaluating ι_{AF} is more problematic:

$$\begin{aligned} \iota_{AF} =& E[A_{i,n}F_{i,n}] \\ =& E[(\mathbf{b}_1'\mathbf{z}_{i,n-1} + \epsilon_{i,n})(\mathbf{b}_3'\mathbf{z}_{i,n-1} + \xi_{i,n})] \\ =& \mathbf{b}_1'E[\mathbf{z}_{i,n-1}\mathbf{z}_{i,n-1}']\mathbf{b}_3 + \text{Cov}(\epsilon_{i,n}, \xi_{i,n}). \end{aligned}$$

Without values for $\mathbf{b}_3$ and $\text{Cov}(\epsilon_{i,n}, \xi_{i,n})$ we cannot estimate ι_{AF}. Boundaries can be put on $\mathbf{b}_3$ from the fact that

$$\begin{aligned} E[P_{i,n}F_{i,n}] = 0 =& E[\omega_{i,n-1}F_{i,n}] = E[\boldsymbol{\beta}'\mathbf{z}_{i,n-1}(\mathbf{b}_3'\mathbf{z}_{i,n-1} + \xi_{i,n-1})'] \\ 0 =& \boldsymbol{\beta}'E[\mathbf{z}_{i,n-1}\mathbf{z}_{i,n-1}']\mathbf{b}_3. \end{aligned}$$

These boundaries are, unfortunately, not very restrictive. There is also no estimate for $\text{Cov}(\epsilon_{i,n}, \xi_{i,n})$. However, bounds can still be placed on the parameters of interest, $\boldsymbol{\beta}$, by estimating $\boldsymbol{\beta}$ for values of $\iota_{AF} \in [-1, 1]$. It shall be seen that the value of ι_{AF} does not have implications for the value of a but only for the values of p and f.

The final assumptions that will be borrowed from the ACE model are that the correlation between the genes of identical twins is $\mu_A = 1$ and between fraternal twins it is $\delta_A = 0.5$. This implies a particular structure on the variance-covariance matrix of $\epsilon_{i,n}$. Particularly, the shock to genes, $\epsilon_{i,n}$ is identical for identical twins, i.e. $\epsilon_{i,n}^{1,mz} = \epsilon_{i,n}^{2,mz}$. Therefore

$$\begin{aligned}
\mu_A =& E[A_{i,n}^{1,mz} A_{i,n}^{2,mz}] \\
=& E[(\mathbf{b}_1'\mathbf{z}_{i,n-1} + \epsilon_{i,n}^{1,mz})(\mathbf{b}_1'\mathbf{z}_{i,n-1} + \epsilon_{i,n}^{2,mz})] \\
=& E[\mathbf{b}_1'\mathbf{z}_{i,n-1}(\mathbf{z}_{i,n-1})'\mathbf{b}_1 + \epsilon_{i,n}^{1,mz}\epsilon_{i,n}^{2,mz} + (\epsilon_{i,n}^{1,mz} + \epsilon_{i,n}^{2,mz})\mathbf{b}_1'\mathbf{z}_{i,n-1}] \\
=& \mathbf{b}_1' E[\mathbf{z}_{i,n-1}(\mathbf{z}_{i,n-1})']\mathbf{b}_1 + \mu_\epsilon \sigma_\epsilon^2 \\
=& \mathbf{b}_1' E[\mathbf{z}_{i,n-1}(\mathbf{z}_{i,n-1})']\mathbf{b}_1 + \sigma_\epsilon^2 = 1 \\
\Rightarrow \sigma_\epsilon^2 =& 1 - \mathbf{b}_1' E[\mathbf{z}_{i,n-1}(\mathbf{z}_{i,n-1})']\mathbf{b}_1.
\end{aligned}$$

This implies

$$\mathbf{b}_1' E[\mathbf{z}_{i,n-1}(\mathbf{z}_{i,n-1})']\mathbf{b}_1 = 0.25$$

and so

$$\sigma_\epsilon^2 = 0.75.$$

Likewise, for fraternal twins,

$$\begin{aligned}
\delta_A =& \mathbf{b}_1' E[\mathbf{z}_{i,n-1}(\mathbf{z}_{i,n-1})']\mathbf{b}_1 + \delta_\epsilon \sigma_\epsilon^2 \\
=& 0.25 + \delta_\epsilon 0.75 = 0.5
\end{aligned}$$

$$\Rightarrow \delta_\epsilon = \frac{1}{3}.$$

Note that $\epsilon_{i,n}$ is orthogonal to $A_{i,n-1}$ and so the correlation between the genes of the parent and child is

$$\begin{aligned}\gamma_A =& E[A_{i,n-1}A_{i,n}]\\ =& E\left[A_{i,n-1}(\mathbf{b}_1'\mathbf{z}_{i,n-1}+\epsilon_{i,n})\right]\\ =& 0.5\,\mathrm{Var}(A_{i,n-1})+E[A_{i,n-1}\epsilon_{i,n}]\\ =& 0.5.\end{aligned}$$

Under these relaxed behavioural genetics assumptions the moment conditions for μ, δ, γ and $\mathrm{Var}(\omega_{i,n})$ are:

$$\mu = a^2 + p^2 + f^2 + 2a\left(p\iota_{AP} + f\iota_{AF}\right), \tag{3.13}$$

$$\delta = 0.5a^2 + p^2 + f^2 + 2a\left(p\iota_{AP} + f\iota_{AF}\right), \tag{3.14}$$

$$\gamma = a\gamma_{\omega A} + p, \tag{3.15}$$

$$\mathrm{Var}(\omega_{i,n}) = 1 = a^2 + p^2 + f^2 + e^2 + 2a\left(p\iota_{AP} + f\iota_{AF}\right), \tag{3.16}$$

where

$$\iota_{AP} = \gamma_{\omega A} = 0.5[a + p\iota_{AP} + f\iota_{AF}]. \tag{3.17}$$

It is immediately obvious that the identification of a^2 and e^2 is unchanged from the standard ACE model. From (3.13) and (3.14) we see that

$$a^2 = 2\left(\mu - \delta\right) \tag{3.18}$$

and from (3.13) and (3.16) that

$$e^2 = 1 - \mu. \tag{3.19}$$

Before moving on to look at identification of p and f in the general case it is instructive to look at identification in two simpler cases. First we examine the case where none of the behavioural genetics assumptions are relaxed - the case where shared environment is uncorrelated with genes, $\iota_{AC} = 0$. Second we examine the case where $\iota_{AF} = 0$. The case where $\iota_{AF} = 0$ allows shared environment to be correlated with genes, $\iota_{AC} \neq 0$, by estimating ι_{AP} in the case that residual family environment is uncorrelated with the genes of the child.

Special Case: $\iota_{AC} = 0$

Proposition 3.1. *If the correlation between common environment and genes is zero,* $\iota_{AC} = 0$ *then:*

$$a = \sqrt{2(\mu - \delta)}, \tag{3.20}$$
$$p = \gamma + \delta - \mu, \tag{3.21}$$
$$f = \sqrt{2\delta - \mu - (\gamma + \delta - \mu)^2}, \tag{3.22}$$
$$e = \sqrt{1 - \mu}. \tag{3.23}$$

Proof:

If $\iota_{AC} = 0$ then it follows that

$$p\iota_{AP} + f\iota_{AF} = c\iota_{AC} = 0.$$

In which case equations (3.13), (3.14), (3.16) and (3.17) reduce to:

$$\mu = a^2 + p^2 + f^2, \tag{3.24}$$

$$\delta = 0.5a^2 + p^2 + f^2, \tag{3.25}$$

$$\text{Var}(\omega_{i,n}) = 1 = a^2 + p^2 + f^2 + e^2, \tag{3.26}$$
$$\iota_{AP} = \frac{a}{2} \tag{3.27}$$

Substituting out ι_{AP} in (3.15), the equation for intergenerational correlations, using (3.27) yields:

$$\gamma = \frac{a^2}{2} + p, \tag{3.28}$$

a^2 and e^2 are identified by:

$$a^2 = 2(\mu - \delta), \tag{3.29}$$
$$e^2 = 1 - \mu. \tag{3.30}$$

Substituting out a^2 in (3.28) and rearranging gives an estimate of p:

$$p = \gamma + \delta - \mu. \tag{3.31}$$

(3.24) and (3.25) imply:

$$2\delta - \mu = p^2 + f^2.$$

Using (3.31) to substitute out p gives the following estimate of f^2:

$$f^2 = 2\delta - \mu - (\gamma + \delta - \mu)^2. \tag{3.32}$$

□

Special Case: $\iota_{AF} = 0$

Proposition 3.2. *If the correlation between residual family environment and genes is zero, $\iota_{AF} = 0$, then the parameters in $\boldsymbol{\beta}$ are uniquely identified as follows:*

$$a = \sqrt{2(\mu - \delta)}, \tag{3.33}$$
$$p = 2 + \gamma - \sqrt{(2 - \gamma)^2 + 4a^2}, \tag{3.34}$$
$$f = \sqrt{2\delta - \mu - (p^2 + 2ap\iota_{AP})}, \tag{3.35}$$
$$e = |\sqrt{1 - \mu}|. \tag{3.36}$$

where

$$\iota_{AP} = \frac{a}{2 - p} \tag{3.37}$$

Proof:

When $\iota_{AF} = 0$ equations (3.13), (3.14), (3.15), (3.16) and (3.17) reduce to:

$$\mu = a^2 + p^2 + f^2 + 2ap\iota_{AP}, \tag{3.38}$$

$$\delta = 0.5a^2 + p^2 + f^2 + 2ap\iota_{AP}, \tag{3.39}$$

$$\gamma = a\gamma_{\omega A} + p, \tag{3.40}$$

$$\mathrm{Var}(\omega_{i,n}) = 1 = a^2 + p^2 + f^2 + e^2 + 2ap\iota_{AP}, \tag{3.41}$$

$$\iota_{AP} = \gamma_{\omega A} 0.5(a + p\iota_{AP}) \tag{3.42}$$

a^2 and e^2 are identified by:

$$a^2 = 2(\mu - \delta), \tag{3.43}$$
$$e^2 = 1 - \mu. \tag{3.44}$$

Rearrange equation (3.40) for $\gamma_{\omega A}$

$$\gamma_{\omega A} = \iota_{AP} = \frac{\gamma - p}{a},$$

then substitute out ι_{AP} in (3.42) and rearrange:

$$p^2 - (2+\gamma)p + 2\gamma - a^2 = 0.$$

The quadratic formula then gives the following two roots for p:

$$p = \frac{2+\gamma) \pm \sqrt{(2-\gamma)^2 + 4a^2}}{2}.$$

Define the upper root as $\bar{p}$ and the lower root as $\underline{p}$. Note that

$$\bar{p} \geq 2.$$

The upper root implies that

$$\gamma_{\omega A} = \frac{\gamma - \bar{p}}{a} < -1.$$

Therefore the upper root is not feasible as it implies a correlation greater than one. Hence

$$p = \underline{p} = \frac{2+\gamma) - \sqrt{(2-\gamma)^2 + 4a^2}}{2}.$$

Substituting out p and ι_{AP} in equations (3.38) and (3.39) implies that

$$2\delta - \mu = f^2 + [\underline{p}^2 + 2a\underline{p}\iota_{AP}]$$

and hence

$$f = \pm\sqrt{2\delta - \mu - [\underline{p}^2 + 2a\underline{p}\iota_{AP}]}.$$

$f \geq 0$ therefore the upper root is the solution for f.

□

General Case: $\iota_{AF} \geq 0$

There is no reason of course to assume that residual family environment is uncorrelated with genes. We are able to identify the parameters of interest for any given value of ι_{AF}. However, we can only prove that we can uniquely identify these parameters for all cases when the correlation between the child's genes and residual family environment is weakly positive (or when $\iota_{AC} = 0$ as shown in the special case above).

Proposition 3.3. *p is uniquely identified for a set of primitives and f as follows:*

$$p = W(f) \equiv \frac{f^2 + af\iota_{AF} + 2\gamma - a^2 - (2\delta - \mu)}{2-\gamma}. \tag{3.45}$$

f is any value of $f \geq 0$ such that $|\iota_{AP}| \leq 1$ and

$$\Phi(f) \equiv f^4 + f^3 b_3 + f^2 b_2 + f b_1 + b_0 = 0. \tag{3.46}$$

Where

$$\begin{aligned}
b_3 =& 2\iota_{AF}\sqrt{2(\mu-\delta)}, \\
b_2 =& 2\left[2\gamma - \mu\right] + 2\left[\mu - \delta\right]\iota_{AF}^2 + \gamma^2 - 4, \\
b_1 =& 2\iota_{AF}\left[4\gamma - \mu - 4\right]\sqrt{2(\mu-\delta)}, \\
b_0 =& \left[2\gamma - \mu\right]\left[2\gamma\left[\gamma - 1\right] - \mu\right] + \left[2\delta - \mu\right]\left[2 - \gamma\right]^2.
\end{aligned}$$

Proof:

$f \geq 0$ by definition. ι_{AP} is a correlation so $|\iota_{AP}| \leq 1$.

Rearrange (3.15) for $\gamma_{\omega A}$:

$$\iota_{AP} = \frac{\gamma - p}{a}. \tag{3.47}$$

Using (3.47) to substitute out ι_{AP} in (3.17) and rearranging yields:

$$Q_p(p) \equiv p^2 - p(2+\gamma) + 2\gamma - a^2 - a\, f \iota_{AF} = 0. \tag{3.48}$$

From equations (3.13) and (3.14)

$$2\delta - \mu = p^2 + f^2 + 2a\left(p\iota_{AP} + f\iota_{AF}\right). \tag{3.49}$$

Again using (3.47) to substitute out ι_{AP} in (3.49) and rearranging

$$Q_f(f) \equiv f^2 + f2a\iota_{AF} + p(2\gamma - p) - (2\delta - \mu) = 0. \tag{3.50}$$

The sum of (3.50) and (3.48) is:

$$Q_p(p) + Q_f(f) = f^2 + af\iota_{AF} - (2-\gamma)p + 2\gamma - a^2 - (2\delta - \mu) = 0 \tag{3.51}$$

This can be rearranged to give p as a function of f:

$$p = W(f) \equiv \frac{f^2 + af\iota_{AF} + 2\gamma - a^2 - (2\delta - \mu)}{2-\gamma}.$$

Which gives equation (3.45) in proposition 3.3. Substituting $W(f)$ into $Q_f(f)$ implies that f is the solution to

$$
\begin{aligned}
0 = Q_f(f) \equiv & f^2 + f2a\iota_{AF} + W(f)(2\gamma - W(f)) - (2\delta - \mu) \\
= & f^4(-1)\frac{1}{(2-\gamma)^2} \\
& + f^3(-2a\iota_{AF})\frac{1}{(2-\gamma)^2} \\
& + f^2\left\{(2-\gamma)^2 - a^2\iota_{AF}^2 + 2\gamma(2-\gamma) - 2\left[2\gamma - a^2 - (2\delta-\mu)\right]\right\}\frac{1}{(2-\gamma)^2} \\
& + f2a\iota_{AF}\left\{(2-\gamma)^2 + \gamma(2-\gamma) - \left[2\gamma - a^2 - (2\delta-\mu)\right]\right\}\frac{1}{(2-\gamma)^2} \\
& + \left\{\left[2\gamma - a^2 - (2\delta-\mu)\right]\left[2\gamma(2-\gamma) - \left[2\gamma - a^2 - (2\delta-\mu)\right]\right] - (2\delta-\mu)(2-\gamma)^2\right\}\frac{1}{(2-\gamma)^2}
\end{aligned}
\tag{3.52}
$$

Substituting in $\sqrt{2(\mu-\delta)}$ for a in (3.52) and multiplying both sides by $\left[-(2-\gamma)^2\right]$ yields:

$$0 = \Phi(f) \equiv f^4 + f^3 b_3 + f^2 b_2 + f b_1 + b_0.$$

□

This implies as many as four pairs of values for f and p that solve the model. However, these roots must also satisfy the following two conditions:

$$f \geq 0 \tag{3.53}$$

$$|\iota_{AP}| = |\gamma_{\omega A}| \leq 1 \tag{3.54}$$

Condition (3.53) follows because f must be weakly positive as F is a latent variable defined as positive only when it has a positive impact on the outcome ω. Condition (3.54) follows, obviously, as $\iota_{AP} = \gamma_{\omega A}$ is a correlation.

It can be shown that so long as the correlation between parental genes and the residual family environment is weakly positive that a solution to (3.46) satisfying conditions (3.53) and (3.54), when it exists, will be unique. As shown in proposition 3.4 below:

Proposition 3.4. *If $\iota_{AF} \geq 0$ there is at most one solution for p and $f \geq 0$ where $|\iota_{AP}| \leq 1$.*

Proof:

When $\iota_{AF} \geq 0$ and $f \geq 0$ then from (3.3.2):

$$\left.\frac{dp}{df}\right|_{\iota_{AF}\geq 0, f\geq 0} = \frac{2f + a\iota_{AF}}{2-\gamma} \geq 0, \tag{3.55}$$

and hence

$$\left.\frac{dQ_f(f)}{df}\right|_{\iota_{AF}\geq 0, f\geq 0} = 2f + 2a\iota_{AF} + 2\frac{dp}{df}(\gamma - p) \geq 0 \text{ where } p \leq \gamma. \tag{3.56}$$

Using the quadratic formula to solve for p in (3.48) yields:

$$p = \frac{2+\gamma \pm \sqrt{(2-\gamma)^2 + 4a(a + f\iota_{AF})}}{2}. \tag{3.57}$$

Call the upper root $\bar{p}$ and the lower root $\underline{p}$. It follows from $f \geq 0$ and $\iota_{AF} \geq 0$ that where roots exist:

$$\bar{p} \geq 2.$$

In conjunction with equation (3.15) this implies that

$$\iota_{AP} = \frac{\gamma - \bar{p}}{a} < -1$$

which contradicts the properties of a correlation. Hence, all feasible values of p are derived from the lower root. It also follows from (3.57) that where roots exist:

$$\underline{p} = \frac{2+\gamma - \sqrt{(2-\gamma)^2 + 4a(a + f\iota_{AF})}}{2} \leq \gamma. \tag{3.58}$$

As the only feasible values of p are less than equal to γ this implies that for feasible solutions of (3.46) where $\iota_{AF} \geq 0$

$$\left.\frac{dQ_f(f)}{df}\right|_{\iota_{AF}\geq 0} \geq 0. \tag{3.59}$$

In which case, in the range of feasible solutions $Phi(f) = 0$ only once.

□

3.4 Counterfactual Analysis

There are four hypothetical experiments that we might want to consider. The first experiment is to equalise all privilege associated with parental income to the mean for generation n such that $P_{i,n} = 0$ for all i. The second experiment is to equalise all privilege associated with parental income to the mean for all generations such that $P_{i,n} = 0$ for all i and n. The third experiment is to equalise the residual family environment to the mean for a single generation n such that $F_{i,n} = 0$ for all i. The fourth experiment is to equalise the residual family environment to the mean for all generations such that $F_{i,n} = 0$ for all i and n.

In order to do this experiment it is useful to have the unnormalised variances and covariances. These are given below:

$$\begin{aligned}\operatorname{Var}(\omega_{i,n}) =& a^2 + p^2 \operatorname{Var}(P_{i,n}) + f^2 \operatorname{Var}(F_{i,n}) + e^2 \\ &+ 2a \left[p \operatorname{Cov}(A_{i,n}, P_{i,n}) + f \operatorname{Cov}(A_{i,n}, F_{i,n})\right]\end{aligned} \tag{3.60}$$

$$\gamma = \frac{E\left[\omega_{i,n-1}\omega_{i,n}\right]}{\sigma_{i,n-1}\sigma_{i,n}} = \frac{1}{\sigma_{i,n-1}\sigma_{i,n}} \left\{a \operatorname{Cov}(\omega_{i,n-1}, A_{i,n}) + p \operatorname{Cov}(\omega_{i,n-1}, P_{i,n})\right\} \tag{3.61}$$

$$\begin{aligned}\operatorname{Cov}(\omega_{i,n-1}, A_{i,n}) =& E\left[\omega_{i,n-1}A_{i,n}\right] = \boldsymbol{\beta}' E\left[z_{i,n-1}z'_{i,n-1}\right] \mathbf{b}_1 \\ =& \frac{1}{2}\left\{a + p \operatorname{Cov}(A_{i,n-1}, P_{i,n-1}) + f \operatorname{Cov}(A_{i,n-1}, F_{i,n-1})\right\}\end{aligned} \tag{3.62}$$

Experiment 1: Equalise all privilege associated with parental income to the mean for generation n such that $P_{i,n} = 0$ for all i. Under this counterfactual define the standard deviation of $\omega_{i,n}$ as σ'_n and the intergenerational elasticity of income between generation $n-1$ and generation n as γ'_n.

As the variance of $P_{i,n}$ is now 0 from equation (3.60) it follows that the variance of income for generation n (only) under this first counterfactual is:

$$(\sigma'_n)^2 = a^2 + f^2 + e^2 + 2af\iota_{AF}. \tag{3.63}$$

Equalisation of privilege occurs only for generation n so the variance of income for generation $n-1$ is unchanged, $\sigma'_{n-1} = 1$. Moreover, while the covariance between the child's genes and environmental privilege falls to zero the covariance between the parents genes and environmental privileges remains the same, ι_{AP}. Consequently, the covariance between parent's income and child's genes is unchanged:

$$\text{Cov}(\omega_{i,n-1}, A_{i,n}) = \iota_{AP} = \frac{a + f\iota_{AF}}{2-p} \tag{3.64}$$

It then follows that the intergenerational elasticity of income under this counterfactual is:

$$\begin{aligned}\gamma'_n =& \frac{1}{\sigma'_{n-1}\sigma'_n}\left\{a\,\text{Cov}(\omega_{i,n-1}, A_{i,n}) + p\,\text{Cov}(\omega_{i,n-1}, P_{i,n})\right\} \\ =& \frac{1}{\sqrt{a^2+f^2+e^2+2af\iota_{AF}}}\left\{\frac{a^2+af\iota_{AF}}{2-p}\right\}.\end{aligned} \tag{3.65}$$

The denominator in the first fraction is the counterfactual standard deviation of $\omega_{i,n}$ multiplied by the counterfactual standard deviation of $\omega_{i,n-1}$. The second fraction, in braces, is the expression we get for $a\,\text{Cov}(\omega_{i,n-1}, A_{i,n})$ from (3.64). While

$$p\,\text{Cov}(\omega_{i,n-1}, P_{i,n}) = 0$$

as $\text{Var}(P_{i,n}) = 0$.

Experiment 2: Equalise all privilege associated with parental income to the mean for all generations such that $P_{i,n} = 0$ for all i and n. Under this counterfactual define the standard deviation of $\omega_{i,n}$ as σ''_n and the intergenerational elasticity of income between generation $n-1$ and generation n as γ''_n.

Under this counterfactual the variance of ω for all generations becomes:

$$(\sigma''_n)^2 = a^2 + f^2 + e^2 + 2af\iota_{AF} \text{ for all } n. \tag{3.66}$$

As the equalisation of privilege occurs for all generations the covariance of parents' genes with their own privilieges is also zero, Hence,

$$\text{Cov}(\omega_{i,n-1}, A_{i,n}) = \frac{a + f\iota_{AF}}{2}. \tag{3.67}$$

It then follows that the intergenerational elasticity of income under this counterfactual is:

$$\begin{aligned}\gamma_n'' =& \frac{1}{\sigma_{n-1}''\sigma_n''}\left\{a\,\mathrm{Cov}(\omega_{i,n-1},A_{i,n}) + p\,\mathrm{Cov}(\omega_{i,n-1},P_{i,n})\right\}\\ =& \frac{1}{a^2+f^2+e^2+2af}\left\{\frac{a^2+af\iota_{AF}}{2}\right\}.\end{aligned} \tag{3.68}$$

The denominator in the first fraction is the counterfactual standard deviation of $\omega_{i,n}$ multiplied by the counterfactual standard deviation of $\omega_{i,n-1}$. The second fraction, in braces, is the expression we get for $a\,\mathrm{Cov}(\omega_{i,n-1},A_{i,n})$ from (3.67). While again,

$$p\,\mathrm{Cov}(\omega_{i,n-1},P_{i,n}) = 0$$

as $\mathrm{Var}(P_{i,n}) = 0$.

Experiment 3: Equalise the residual family environment the mean for a single generation n such that $F_{i,n} = 0$ for all i. Under this counterfactual define the standard deviation of $\omega_{i,n}$ as σ_n^* and the intergenerational elasticity of income between generation $n-1$ and generation n as γ_n^*.

As the variance of $F_{i,n}$ is now 0 from equation (3.60) it follows that the variance of income for generation n (only) under this first counterfactual is:

$$(\sigma_n^*)^2 = a^2 + p^2 + e^2 + 2ap\iota_{AP}. \tag{3.69}$$

Residual family environment is unchanged in the parental generation and so the covariance between the child of generation n's genes and parental income (or equivalently privilege) is unchanged. Hence, we can use ι_{AP}/γ_{AW} as estimated prior to the counterfactual analysis. This implies that the intergenerational elasticity of income rises to:

$$\gamma^* = \frac{1}{\sigma_n^*}\gamma \tag{3.70}$$

Experiment 4: Equalise the residual family environment the mean for all generations such that $F_{i,n} = 0$ for all i and n. Under this counterfactual define the standard deviation of $\omega_{i,n}$ as σ_n^{**} and the intergenerational elasticity of income between generation $n-1$ and generation n as γ_n^{**}.

Consider first that as $\mathrm{Var}(F_{i,n}) = 0$ for all generations that $\mathrm{Cov}(A_{i,n-1},F_{i,n-1}) = 0$. Hence:

$$\mathrm{Cov}(A_n, P_n) = \mathrm{Cov}(\omega_{i,n-1}, A_{i,n}) = \frac{1}{2}\left[a + p\,\mathrm{Cov}(\omega_{i,n-2}, A_{i,n-1})\right]. \tag{3.71}$$

From the stationarity condition

$$\mathrm{Cov}(\omega_{i,n-1}, A_{i,n}) = \mathrm{Cov}(\omega_{i,n-2}, A_{i,n-1}) \text{ for all } n$$

and so

$$\mathrm{Cov}(A_n, P_n) = \mathrm{Cov}(\omega_{i,n-1}, A_{i,n}) = \frac{a}{2-p} \text{ for all } n. \tag{3.72}$$

This implies then that the counterfactual variance of income is

$$(\sigma_n^{**})^2 = a^2 + p^2 + e^2 + 2p\frac{a^2}{2-p}. \tag{3.73}$$

and the counterfactual intergenerational elasticity of income is

$$\gamma^{**} = \frac{1}{(\sigma_n^{**})^2}\left\{\frac{a^2}{2-p} + p(\sigma_n^{**})^2\right\}. \tag{3.74}$$

3.5 An Application to Taubman (1976) and Lee and Solon (2009)

Taubman (1976) uses American male twins born between 1917 and 1927 data from the NAS-NRC twins who did military service during World War 2. Taubman's study does not, however, report intergenerational elasticities of income. The closest estimates for the intergenerational elasticities of this cohort can be found in Lee and Solon (2009). In these two papers the following elasticities between the incomes of identical twins, fraternal twins and between fathers and sons are:

$$\mu = 0.54, \qquad \delta = 0.30, \qquad \text{and} \qquad \gamma = 0.34. \tag{3.75}$$

From these results the estimates of a^2 and e^2 are:

$$\hat{a}^2 = 2(\mu - \delta) = 0.48 \tag{3.76}$$

$$\hat{e}^2 = 1 - \mu = 0.46 \tag{3.77}$$

Note, these estimates do not depend on the correlation of genes and residual shared environment, ι_{AF}. However, the estimates of p, f and ι_{AP} do depend on the value of ι_{AF}. Estimates of the parameters of interest are given for values of ι_{AF} over the interval $[0,1]$.[5]

Table 1: Parameter Estimates using twin data from Taubman (1976) and intergenerational data from Lee and Solon (2009).

	$\iota_{AF}=0$	$\iota_{AF}=0.25$	$\iota_{AF}=0.5$	$\iota_{AF}=1$
p	0.089	0.087	0.087	0.087
f	0.086	0.023	0.012	0.0061
ι_{AP}	0.36	0.37	0.37	0.37
$a\iota_{AP}$	0.251	0.253	0.253	0.253

The environmental impact of parental income and its correlates are quite insensitive to variation in the correlation between the residual shared environment and genes. Likewise, the correlation between parental income and the child's genes is similarly insensitive to variation in ι_{AF}.

The first counterfactual experiment of increasing the environmental privileges associated with paternal income by one percent is p. This is close to a one tenth of a percentage point increase for all ι_{AF}.

The second counterfactual experiment is to equalise, to the population mean, the environmental privilege associated with paternal income:

Table 2: The effects of equalising environmental privilege associated with paternal income.

	$\iota_{AF}=0$	$\iota_{AF}=0.25$	$\iota_{AF}=0.5$	$\iota_{AF}=1$
Fall in Inequality	0.052	0.052	0.052	0.052
Fall in IGE	0.1	0.098	0.098	0.098
Percentage Fall in IGE	0.294	0.288	0.288	0.288

These results are also relatively invariant to ι_{AF}. This should be unsurprising as the effect of removing privilege on inequality and intergenerational mobility is determined by p and ι_{AP}. Inequality, as measured by income variance, falls by about 5 percent and the intergenerational elasticity of income falls by close to 30%.

The third counterfactual experiment in which we are interested is the effect of equalising the residual shared environment to the mean. It might be expected that the results of this

[5] The reported results are based on the third root of the quartic equation (**??**). All other roots imply either a negative value for f or that $|\iota_{AP}|>1$.

experiment are more sensitive to ι_{AF} as the effect of the residual family environment varies greatly with ι_{AF}, decreasing more than 10 fold when moving from $\iota_{AF} = 0$ to $\iota_{AF} = 1$. In the case of moving someone in the distribution of residual common environment by one percent the effect is between one tenth of a percent and 1 two hundredth of a percent. Note that the effect drops quite quickly with ι_{AF}, a move from $\iota_{AF} = 0$ to $\iota_{AF} = 0.25$ implies a four fold drop in the impact of the residual common environment.

However, despite the range in values for f, we get the following results:

Table 3: The effects of equalising residual family environment.

	$\iota_{AF} = 0$	$\iota_{AF} = 0.25$	$\iota_{AF} = 0.5$	$\iota_{AF} = 1$
Fall in Inequality	0.007	0.008	0.008	0.008
Fall in IGE	0	0.002	0.002	0.002
Percentage Fall in IGE	0	0.006	0.006	0.006

As with the experiment of equalising privilege the effect of equalising the residual shared environment to the mean varies little with ι_{AF}. The impact on inequality is to reduce inequality by a little less than one percent. Intergenerational elasticity of income falls by as much as six tenths of a percent.

To the extent that inequality is caused by family environment this analysis suggests that it is mostly due to those privileges associated with parental income. Likewise, the intergenerational persistence of earnings, to the extent it is caused by family environment, is caused by those privileges associated with parental income.

The difference in the impact of the two experiments are relatively constant in ι_{AF}. However, in general the extent to which the paternal income measure captures shared environment varies greatly in ι_{AF}. As the impact of residual common environment decreases with ι_{AF} this implies the proportion of the common environment explained by the privilege associated with parental income increases. This can be seen below:

Table 4: The importance of paternal income in explaining family environment.

	$\iota_{AF} = 0$	$\iota_{AF} = 0.25$	$\iota_{AF} = 0.5$	$\iota_{AF} = 1$
$\frac{p^2}{p^2 + f^2}$	0.513	0.936	0.981	0.995

Paternal income captures at least fifty percent of common environment. Note, however, for correlations greater than or equal to 0.25 paternal income captures more than ninety percent of common environment.

3.6 Conclusion

The methodology developed in this paper allows us use twin data in combination with data on parents to answer the following questions:

1. How much of intergenerational mobility is due to the privileges associated with parental income?

2. What is the effect of variation in the privileges associated with parental income on inequality?

3. How much of family environment is explained by a measure of parental income?

To answer these questions this method requires a subset of the assumptions used within classical twin studies. Note, that the ACE model approach does not give answers to any of these questions.

In the speculative application of this methodology to existing data on American twins and the intergenerational elasticity of income between fathers and sons I get the following answers to these three questions: 1) About thirty percent of the intergenerational elasticity of income is due to the privileges associated with parental income; 2) About 5 percent of income variance is caused by privileges associated with parental income; and 3) Between fifty and ninety nine per cent of family environment is explained by a measure of parental income.

These results must, however, be treated with extreme scepticism as the estimate of intergenerational elasticity is for children born in 1952 while the twin correlations are for children born between 1917 and 1927. That the intergenerational elasticity may be too high or too low is a source of bias for the estimate of p, f and ι_{AP}. Similarly, the different methods for collection of the two sources of data could lead to bias. The solution to this problem, of course, is better data. The next logical step it to take this method to a data set where there are measurements of both the twins and their parents.

As stated earlier there are also several issues with the the ACE model that are also problems for the method developed in this paper. First, there are several covariance restrictions that are carried over from the ACE model. Second, the assumption of additive genetics. This is an assumption that the interaction effects between an individuals genetic polymorphisms are negligible. There is some evidence to suggest that this issue is not particularly problematic (Neale and Cardon, 1992) but the issue is far from settled. Third, and perhaps most seriously, is the assumptions of additive separability and linearity. Heckman (2007) citing Rutter (2006) criticises the notion that nature and nurture can be neatly decomposed into two separate effects.

This paper allows us to deal with one of these issues, the assumption of family environment being uncorrelated with genes. More importantly the method in this paper allows us to interpret the twin study results in a way that is meaningful for the growing debate over inequality and intergenerational mobility. This said, there is clearly much still to be done if we are to understand what data on twins really means for the causes of inequality and intergenerational mobility.

Bibliography

Abreu, D. and M. Brunnermeier (2003). Bubbles and Crashes. *Econometrica 71*(1), 173–204.

Acemoglu, D. and M. O. Jackson (2011). History, Expectations, and Leadership in Evolution of Social Norms. *SSRN Electronic Journal*.

Akerlof, G. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The quarterly journal of economics*.

Banerjee, A. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). A theory of fads, fashion, custom and cultural change as informational cascades. *Journal of political Economy*.

Björklund, A. and M. Jäntti (2009). Influences of Nature and Nurture on Earnings Variation: A Report on a Study of Various Sibling Types in Sweden. In S. Bowles, H. Gintis, and M. O. Groves (Eds.), *Unequal Chances. Family Background and Economic Success*, pp. 145–164. Princeton University Press.

Blanden, J., P. Gregg, and L. Macmillan (2007). Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education. *The Economic Journal 117*(519).

Boomsma, D., A. Busjahn, and L. Peltonen (2002, November). Classical twin studies and beyond. *Nature reviews. Genetics 3*(11), 872–82.

Bound, J. and G. Solon (1999, April). Double trouble: on the value of twins-based estimation of the return to schooling. *Economics of Education Review 18*(2), 169–182.

Bowles, S. and H. Gintis (2002). The Inheritance of Inequality. *The Journal of Economic Perspectives 16*(3), 3–30.

Bowles, S., H. Gintis, and M. Osborne (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature XXXIX* (December), 1137–1176.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014). Is The United States Still A Land of Opportunity? Recent Trends in Intergenerational Mobility.

Corak, M. (2011). Inequality from generation to generation: the United States in comparison. In R. Rycroft (Ed.), *The Economics of Inequality, Poverty, and Discrimination in the 21st Century*, pp. 107–123. Praeger Publishers Inc.

Dow, J. (2004). Is Liquidity SelfFulfilling?*. *The Journal of Business 77* (4), 895–908.

Eisfeldt, A. L. (2004). Endogenous liquidity in asset markets. *The Journal of Finance LIX* (1).

Fox, H. E. and M. V. Erdmann (2000). Fish yields from blast fishing in indonesia. *Coral Reefs 19* (2), 114–114.

Fudenberg, D. and J. Tirole (1991). *Game Theory.* Cambridge MA: MIT Press.

Goldberger, A. S. (1979). Heritability. *Economica 46* (184), 327–347.

Guerrieri, V. and R. Shimer (2013). Dynamic Adverse Selection: A Theory of Illiquidity, Fire Sales, and Flight to Quality.

Heckman, J. J. (2007, August). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences of the United States of America 104* (33), 13250–5.

Heckman, J. J. and S. Mosso (2014). The Economics of Human Development and Social Mobility.

Hendel, I. and A. Lizzeri (1999). Adverse Selection in Durable Goods Markets. *American Economic Review 89* (5), 1097–1115.

Hendel, I., A. Lizzeri, and M. Siniscalchi (2005, April). Efficient Sorting in a Dynamic Adverse-Selection Model. *Review of Economic Studies 72* (2), 467–497.

Hermalin, B. E. (1998). Toward an economic theory of leadership: Leading by example. *The American Economic Review 88* (5), 1188–1206.

Herrnstein, R. and C. Murray (1994). *Bell curve: Intelligence and class structure in American life.* Simon and Schuster.

Isacsson, G. (1999, November). Estimates of the return to schooling in Sweden from a large sample of twins. *Labour Economics 6* (4), 471–489.

Jantti, M., B. Bratsberg, K. Rø ed, and O. Raaum (2006). American exeptionalism in a new light: A comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States.

Kandori, M. (1992). Social norms and community enforcement. *The Review of Economic Studies*.

Kelling, G. L. and J. Q. Wilson (1982). Broken windows. *Atlantic Monthly 249*(3), 29–38.

Kropotkin, P. A. (1887). *In Russian and French Prisons.* Ward and Downey.

Krueger, A. (2012). The rise and consequences of inequality. Technical report, Center for American Progress.

Krueger, A. and O. Ashenfelter (1994). Estimates of the economic return to schooling from a new sample of twins. *The American Economic Review 84*(5), 1157–1173.

Lee, C. and G. Solon (2009). Trends in Intergenerational Income Mobility. *The Review of Economics and Statistics* (February).

Mailath, G. J. and L. Samuelson (2006). *Repeated games and reputations: long-run relationships.* Oxford University Press, USA.

Manski, C. F. (2011, November). Genes, Eyeglasses, and Social Policy. *Journal of Economic Perspectives 25*(4), 83–94.

Merriman, C. (1924). The intellectual resemblance of twins. *Psychological Monographs: General and Applied xxxi*(5).

Neale, M. and L. Cardon (1992). *Methodology for genetic studies of twins and families.* Springer.

Rutter, M. (2006). *Genes and behavior: Nature-nurture interplay explained.* Wiley-Blackwell.

Shiller, R. (1980). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *The American Economic Review 71*(3), 421–436.

Shiller, R. (1992). *Market volatility.* MIT Press.

Taubman, P. (1976). The Determinants of Earnings: Genetics, Family, and Other Environments; A Study of White Male Twins. *The American Economic Review*.

Turkheimer, E. (2000). Three Laws of Behavior Genetics and What They Mean. *Current Directions in Psychological Science 9*(5), 160–164.