

Centre interuniversitaire de recherche
en économie quantitative

CIREQ

Cahier 04-2011

*Dynamic Linear Economies with
Social Interactions*

Onur ÖZGÜR and
Alberto BISIN



Le **Centre interuniversitaire de recherche en économie quantitative (CIREQ)** regroupe des chercheurs dans les domaines de l'économétrie, la théorie de la décision, la macroéconomie et les marchés financiers, la microéconomie appliquée et l'économie expérimentale ainsi que l'économie de l'environnement et des ressources naturelles. Ils proviennent principalement des universités de Montréal, McGill et Concordia. Le CIREQ offre un milieu dynamique de recherche en économie quantitative grâce au grand nombre d'activités qu'il organise (séminaires, ateliers, colloques) et de collaborateurs qu'il reçoit chaque année.

*The **Center for Interuniversity Research in Quantitative Economics (CIREQ)** regroups researchers in the fields of econometrics, decision theory, macroeconomics and financial markets, applied microeconomics and experimental economics, and environmental and natural resources economics. They come mainly from the Université de Montréal, McGill University and Concordia University. CIREQ offers a dynamic environment of research in quantitative economics thanks to the large number of activities that it organizes (seminars, workshops, conferences) and to the visitors it receives every year.*

Cahier 04-2011

Dynamic Linear Economies with Social Interactions

Onur ÖZGÜR and Alberto BISIN

Ce cahier a également été publié par le Département de sciences économiques de l'Université de Montréal sous le numéro (2011-03).

This working paper was also published by the Department of Economics of the University of Montreal under number (2011-03).

Dépôt légal - Bibliothèque nationale du Canada, 2011, ISSN 0821-4441

Dépôt légal - Bibliothèque et Archives nationales du Québec, 2011

ISBN-13 : 978-2-89382-613-4

Dynamic linear economies with social interactions ^{*}

Onur Özgür

Université de Montréal[†]
CIREQ, CIRANO

Alberto Bisin

New York University[‡]
NBER and CIREQ

February 15, 2011

Abstract

Social interactions arguably provide a rationale for several important phenomena, from smoking and other risky behavior in teens to e.g., peer effects in school performance. We study *social interactions* in *dynamic* economies. For these economies, we provide existence (Markov Perfect Equilibrium in pure strategies), ergodicity, and welfare results. Also, we characterize equilibria in terms of agents' policy function, spatial equilibrium correlations and social multiplier effects, depending on the nature of interactions. Most importantly, we study formally the issue of the identification of social interactions, with special emphasis on the restrictions imposed by dynamic equilibrium conditions.

Journal of Economic Literature Classification Numbers: C31, C62, C72, C73, Z13.

Keywords: Conditional covariance stationarity, conformity, ergodicity, habits, identification, local and global interactions, Markov perfect equilibrium, social interactions, social norms, social processes, social status, spatial correlations.

^{*}This paper is part of the Polarization and Conflict Project CIT-2-CT-2004-506084 funded by the European Commission-DG Research Sixth Framework Programme. We are grateful to Bernard Salanié for some insightful comments early on. We also thank Massimiliano Amarante, Jess Benhabib, Ken Binmore, Michele Boldrin, Yann Bramoullé, Pierre-André Chiappori, In-Koo Cho, Aureo De Paula, Itzhak Gilboa, Bryan Graham, Ali Hortaçsu, Frédéric Koessler, Takashi Kunimoto, Justin Leroux, David Levine, Bob Lucas, Andy McLennan, Debraj Ray, Larry Samuelson, Manuel Santos, Tom Sargent, José Scheinkman, Karl Schlag, David Schmeidler, Paolo Siconolfi, Yves Sprumont, Bruno Strulovici, Jean Tirole, and many seminar participants. Part of this research was done while Özgür was visiting the Economics Department at the Université Laval. Thanks to Yann Bramoullé and Bernard Fortin for organizing the visit. Özgür is grateful for financial support to “La Chaire du Canada en Économie des Politiques Sociales et des Ressources Humaines” at Université Laval, CIREQ, CIRPÉE, and FQRSC.

[†]Department of Economics; e-mail: onur.ozgur@umontreal.ca

[‡]Department of Economics; e-mail: alberto.bisin@nyu.edu

1 Introduction

Agents interact in markets as well as socially, that is, in the various socioeconomic groups they belong to. Models of social interactions are designed to capture in a simple abstract way socioeconomic environments in which markets do not mediate all of agents' choices. In such environments agents' choices are determined by their preferences as well as by their ability to interact with others, on their position in a predetermined network of relationships, e.g., a family, a peer group, or more generally any socioeconomic group.¹

Social interactions arguably provide a rationale for several important phenomena, Peer effects, in particular, have been indicated as one of the main empirical determinants of risky behavior in adolescents.² Relatedly, peer effects have been studied in connection with education outcomes,³ obesity,⁴ friendship and sex,⁵ labor market referrals,⁶ neighborhood and employment segregation,⁷ criminal activity,⁸ and several other socioeconomic phenomena.⁹

The large majority of the existing models of social interactions are static; or, when dynamic models of social interactions are studied, it is typically assumed that agents are myopic and their choices are subject to particular behavioral assumptions.¹⁰ In this paper, we contribute

¹The integration of models of social interactions within economic theory is an active and interesting area of research. See the recent *Handbook of Social Economics*, Benhabib, Bisin, and Jackson (2010).

²See e.g., Ali and Dwyer (2009), Axtell et al. (2006), Bauman and Ennett (1996), Bifulco et al. (2009), Chaloupka and Warner (2000), Clark and Loheac (2007), Cook and Moore (2000), Cutler and Glaeser (2007), DeCicca et al. (2008), Evans et al. (1992), Fletcher (2009), Gaviria and Raphael (2001), Gilleskie and Strumpf (2005), Gilleskie and Zhang (2010), Jones (1984), Kobus (2003), Krauth (2005, 2006), Kremer and Levy (2008), Krosnick and Judd (1982), Lewitt et al. (1981), Lundborg (2006), Nakajima (2007), Norton et al. (1998), Powell et al. (2003), Sacerdote (2001), Soetevent and Kooreman (2007), Tyas and Pederson (1998), Wang et al. (1995, 2000).

³Altonji et al (2005), Ammermuller and Pischke (2009), Bénabou (1996), Borjas (1995), Boozer and Cacciola (2001), Carrell et al. (2009), De Giorgi et al. (2009), Evans et al. (1992), Gaviria and Raphael (2001), Hoxby (2000a, 2000b), Soetevent and Kooreman (2007), Zimmerman (2003).

⁴Burke and Heiland (2007), Christakis and Fowler (2007).

⁵Akerlof et al. (1996), Bearman et al. (2004), Bramoullé et al. (2009), Broadhead et al. (1998), Cipollone and Rosolia (2007), Conti et al. (2009), Currarini et al. (2009), Kandel (1978), Leider et al. (2007), Mihaly (2007), Moody (2001).

⁶Bayer et al. (2008), Bjorn and Vuong (1985), Calvo-Armengol and Jackson (2004), Conley and Topa (2002, 2007), De Giorgi et al. (2009), Goldin and Katz (2002), Granovetter (1973, 1995), Grodner and Kniesner (2007), Ichino and Falk (2006), Ioannides and Datcher Loury (2004), Moro (2003), Topa (2001), Weinberg et al. (2004).

⁷Aizer and Currie (2004), Bénabou (1993), Case and Katz (1991), Crane (1991), Durlauf (1996, 2004), Goering and Feins (1997), Hoff and Sen (2005), Ioannides and Topa (2009), Ioannides and Zabel (2008), Ioannides and Zarella (2009), Katz et al. (2001), Ludwig et al. (2001), Mobius (2000), Rosenbaum (1995), Schelling (1971, 1972).

⁸Calvo Armengol et al. (2009), Glaeser et al. (1996), Kling et al. (2005), Ludwig et al. (2001).

⁹See Bisin et al. (2010), Glaeser and Scheinkman (2001), Moffitt (2001) for surveys.

¹⁰Exceptions include an example on female labor force participation in Glaeser and Scheinkman (2001), Binder and Pesaran (2001) on life-cycle consumption under social Interactions, Blume (2003) on social stigma, Brock and Durlauf (2010) and de Paula (2009) on duration models, and the theoretical analysis of Bisin, Horst, and Özgür

to this literature by studying social interactions in *dynamic* economies. We focus our attention on *linear economies*, in which each agent's preferences are quadratic. Dynamic linear models of course have appealing analytical properties. Hansen and Sargent (2004) study this class of models systematically, exploiting the tractability of linear control methods and matrix Riccati equations. While the class of economies we study in this paper allows however for a countable number of heterogeneous agents and an infinite horizon, giving rise to infinite dimensional systems, some tractability is maintained. Furthermore, in the class of economies we study agents display preferences for *conformity*, that is, preferences which incorporate the desire to conform to the choices of agents in a reference group.

More specifically, each agent's preferences are hit by random preference shocks over time. Each agent interacts with agents in his social reference group, in the sense that each agent's instantaneous preferences depend on the current choices of agents in his social reference group, as a direct externality. Each agent's instantaneous preferences also depend on the agent's own previous choice, representing the inherent costs to dynamic behavioural changes due e.g., to habits. When agents' reference groups overlap, each agent's *optimal* choice depends on all the other agents's previous choices and current preference shocks, as long as they are observable. We allow for complete and incomplete information with respect to preference shocks. Requiring that the social and informational structure of each agent satisfy a symmetry condition, we restrict our analysis to *symmetric Markov perfect equilibria*. Agents' choices at equilibrium are determined by *linear* policy (best reply) functions. More specifically, e.g., in infinite-horizon economies, a symmetric Markov perfect equilibrium is represented by a symmetric policy function, for each arbitrary agent $a \in \mathbb{A}$, a countable set, which maps the agent's choice at time t , x_t^a , linearly in each agent's past choices, x_{t-1}^{a+b} , in each agent's contemporaneous idiosyncratic preference shock, θ_t^{a+b} , and in the mean preference shock, $\bar{\theta}$:

$$x_t^a = \sum_{b \in \mathbb{A}} c^b x_{t-1}^{a+b} + \sum_{b \in \mathbb{A}} d^b \theta_t^{a+b} + e \bar{\theta}$$

For these economies, we provide some fundamental theoretical results: (Markov perfect) equilibria exist (for finite economies they are unique) and they induce an ergodic stochastic process over the equilibrium configuration of actions. Furthermore, a stationary ergodic distribution exists. We also derive a recursive algorithm to compute equilibria. The proof of the existence theorem, in particular, requires some subtle arguments. In fact, standard variational arguments require to bound the marginal effect of any infinitesimal change dx^a on the agent's value function. But in the class of economies we study, the Envelope theorem (as e.g., in Benveniste and Scheinkman (1979)) is not sufficient for this purpose, as dx^a affects agent a 's value function directly and indirectly, through its effects on all agents $b \in \mathbb{A} \setminus a$'s choices, which in turn affect agent a 's value function. The marginal effect of any infinitesimal change dx^a is then an infinite sum of

(2006).

endogenous terms. In our economy, however, we can exploit the linearity of policy functions to represent a symmetric MPE by a fixed point of a recursive map which can be directly studied.

Exploiting the linear structure of our economies we can study equilibria in some detail, characterizing the parameters of the policy function as well as a fundamental statistical property of equilibrium, the cross-sectional auto-correlation of actions. In turn we obtain a series of results regarding the welfare properties of equilibrium and various comparative dynamics exercises of interest. First of all, we show that, since social interactions are modelled in this paper as a preference externality, equilibria will not be efficient in general. We also characterize the form of the inefficiency: at equilibrium each agent's policy function weights too heavily the agent's own preference shock and previous action and not enough the other agents'. The comparative dynamics exercises illustrate e.g., the equilibrium effects of the strength of social interactions and of the social and informational structure of the economy.

Finally, we exploit our characterization results of the equilibria to address generally the issue of identification of social interactions in our context, with population data. While the empirical literature has often interpreted a significant high correlation of socioeconomic choices across agents, e.g., peers, as evidence of social interactions, in the form e.g., of preferences for conformity, it is well known at least since the work of Manski (1993) that the empirical study of social interactions is plagued by subtle identification problems. Intuitively, in our economy for instance, the spatial correlation of actions at equilibrium can be due to social interactions or to the spatial correlation of preference shocks. More formally, take two agents, e.g., agent a and agent b . A positive correlation between x_t^a and x_t^b could be due to e.g., preference for conformity. But the positive correlation between x_t^a and x_t^b could also be due to a positive correlation between θ_t^a and θ_t^b . In this last case, preferences for conformity and social interactions would play no role in the correlation of actions at equilibrium. Rather, such correlation would be due to the fact that agents have correlated preferences. Correlated preferences could generally be due to some sort of assortative matching or positive selection, which induce agents with correlated preferences to interact socially.

In the context of our economy, we ask whether the restrictions implied by the dynamic equilibrium analysis help identify social interactions and distinguish them from correlated preferences. We show that the answer is in fact affirmative, but only if the economy is non-stationary, in a precise sense. To illustrate our results, consider for instance the issue of peer effects in adolescents' substance use. Suppose the econometrician observes the behavior of a population of students in a school over time (at different grades). A significant high correlation of socioeconomic choices across students in the school could be due to selection in the endogenous composition of the school in terms of unobserved (to the econometrician) correlated characteristics of the agents. Any significant variation in students' behavior through time (grades) must however be due to social interactions. A student whose choice is affected by the choices of his school peers will in fact

rationally anticipate how much longer he will interact with them. In particular, his propensity to conform to his peers' actions will tend to decrease over time (grades) and will be the lowest in the final years in the school. This non-stationarity of each student's behavior at equilibrium is the key to the identification of social interaction in our class of economies.¹¹

The simplicity of linear models allows us to extend our analysis in several directions which are important in applications and empirical work. This is the case, for instance of general (including asymmetric) neighborhood network structures for social interactions. But our analysis extends also to general stochastic processes for preference shocks and to the addition of global interactions. One particular form of global interactions occurs when each agent's preferences depend on an average of actions of all other agents in the population, e.g. Brock and Durlauf (2001a), and Glaeser and Scheinkman (2003). This is the case, for instance, if agents have preferences for *social status*. More generally, global interactions could capture preferences to adhere to aggregate norms of behavior, such as specific group cultures, or other externalities as well as price effects. Finally, and perhaps most importantly, we extend our analysis to encompass a richer structure of dynamic dependence of agents' actions at equilibrium. In particular we study an economy in which agents' past behavior is aggregated through an accumulated stock variable which carries habit persistence, which can be directly applied e.g., to the issue of teenage substance addiction due to peer pressure at school. With respect to the addiction literature, as e.g., Becker and Murphy (1988), we model the dynamics of addiction considering peer effects not only in a single-person decision problem, but rather as an equilibrium effect allowing for the intertemporal feedback channel between agents across social space and through time.¹² In this context we show that in equilibrium each agent's choice depends on the stock of his neighbors' actions, on their long-term behavioral patterns rather than just on their previous period actions. Also, in non-stationary economies, as the final period approaches, each agent assigns higher weights to his own stock, giving rise to an initiation-addiction behavioral pattern at equilibrium which is consistent with observation, e.g., in Cutler and Glaeser (2007) and DeCicca, Kenkel, and Mathios (2008).

2 Dynamic economies with social interactions

While we develop most of our analysis in the context of linear models, it is useful to set up the general model first, as we do in this section, to be as clear and specific as possible regarding the assumptions we impose on the economy we study.

Time is discrete and is denoted by $t = 1, \dots, T$. We allow both for infinite economies ($T = \infty$) and economies with an end period ($T < \infty$). A typical economy is populated by a countable

¹¹This pattern of behavior appears consistent with the peer effects study of Hoxby (2000a,b).

¹²See also Becker, Grossman, and Murphy (1994), Boyer (1978, 1983), Gul and Pesendorfer (2007), Gruber and Koszegi (2001), Iannaccone (1986); see Rozen (2010) for theoretical foundations for intrinsic linear habit formation; see also Elster (1999) and Elster and Skog (1999) for surveys.

number of *agents* $a \in \mathbb{A}$.¹³ Each agent lives for the duration of the economy. At the beginning of each period t , agent a 's random preference *type* θ_t^a is drawn from Θ , a compact subset of a finite dimensional Euclidean space $\mathbb{R}^n$. The random variables θ_t^a are independently and identically distributed across time and agents with probability law ν . We assume, with no loss of generality, that the random variable $\theta_t := (\theta_t^a)_{a \in \mathbb{A}}$ is defined, for all t , on the canonical probability space $(\Theta, \mathcal{F}, \mathbb{P})$, where $\Theta := \{(\theta^a)_{a \in \mathbb{A}} : \theta^a \in \Theta\}$. At each period t , agent $a \in \mathbb{A}$ chooses an *action* x_t^a from the set X , a compact subset of a finite dimensional Euclidean space $\mathbb{R}^p$. Let $\mathbf{X} := \{x = (x^a)_{a \in \mathbb{A}} : x^a \in X\}$ be the space of individual action profiles.

Each agent $a \in \mathbb{A}$ *interacts* with agents in the set $N(a)$, a nonempty subset of the set of agents $\mathbb{A}$, which abstractly represents agent a 's social reference group. The map $\mathbb{A} : N \rightarrow 2^{\mathbb{A}}$ is referred to as a *neighbourhood correspondence* and is assumed exogenous. Agent a 's instantaneous preferences depend on the current choices of agents in his reference group, $\{x_t^b\}_{b \in N(a)}$, representing social interactions as direct preference externalities. Agent a 's instantaneous preferences also depend on the agent's own previous choice, x_{t-1}^a , representing inherent costs to dynamic behavioural changes due e.g., to habits. In summary, agent a 's instantaneous preferences at time t are represented by a continuous utility function

$$\left(x_{t-1}^a, x_t^a, \{x_t^b\}_{b \in N(a)}, \theta_t^a\right) \mapsto u\left(x_{t-1}^a, x_t^a, \{x_t^b\}_{b \in N(a)}, \theta_t^a\right)$$

Agents discount expected future utilities using the common stationary discount factor $\beta \in (0, 1)$.

The economy has an exogenous initial configuration $x_0 \in \mathbf{X}$. Let $x^{t-1} = (x_0, x_1, \dots, x_{t-1})$ and $\theta^{t-1} = (\theta_1, \dots, \theta_{t-1})$ be the $(t-1)$ -period choices and type realizations. Before each agent's time t choice, x^{t-1} is observed by all agents and the current value of the random variable θ_t realizes. Agent $a \in \mathbb{A}$ observes only the part $I_a \theta_t := \{\theta_t^b : b \in I(a)\}$, where $I(a) \subset \mathbb{A}$ is his *information set*. Similarly, let $I_a \theta^{t-1} = (I_a \theta_1, \dots, I_a \theta_{t-1})$. We study both economies with *complete information*, $I(a) = \mathbb{A}$, and economies with *incomplete information*, $I(a) \subsetneq \mathbb{A}$. After each agent's time t choice, $x_t = (x_t^b)_{b \in \mathbb{A}} \in \mathbf{X}$ becomes common knowledge and the economy moves to time $t+1$.

A *strategy* for an agent a is a sequence of measurable functions $x^a = (x_t^a)$, where for each t , $x_t^a : \mathbf{X}^t \times (\Theta^{I(a)})^t \rightarrow X$. Agents' strategies along with the probability law for types induce a stochastic process over future configuration paths. Each agent $a \in \mathbb{A}$'s objective is to choose x^a to maximize

$$E \left[\sum_{t=1}^T \beta^{t-1} u \left(x_{t-1}^a, x_t^a, \{x_t^b\}_{b \in N(a)}, \theta_t^a \right) \mid (x_0, \theta_1) \right] \quad (1)$$

given the strategies of other agents and given $(x_0, I_a \theta_1) \in \mathbf{X} \times \Theta^{I(a)}$.

¹³We study an economy populated by a countably infinite number of agents where $\mathbb{A} := \mathbb{Z}$, but our analysis applies to economies with a finite number of agents.

We require that the social and informational structure satisfies the following symmetry restrictions:¹⁴

1. For all $a, b \in \mathbb{A}$, $N(b) = R^{b-a}N(a)$, where R^{b-a} is the canonical *shift* operator in the direction $b - a$.¹⁵
2. For all $a, b \in \mathbb{A}$, $I(b) = R^{b-a}I(a)$.

We restrict our analysis to *symmetric Markov perfect equilibria*. Agents' strategies are *Markovian* if after any $t - 1$ -period history (x^{t-1}, θ^t) , they depend only on the previous period configuration x_{t-1} and the current type realizations θ_t . Because of symmetry, it is thus enough to analyze the optimization problem relative to a single reference agent, say agent $0 \in \mathbf{A}$. Thus, we assume that the optimal choice of any economic agent $b \in \mathbb{A}$ is determined by a continuous choice function $g : \mathbf{X} \times \Theta^{I(0)} \times \{1, \dots, T\} \rightarrow X$ such that for all $t = 1, \dots, T$ and after any history $(x^{t-1}, \theta^t) \in \mathbf{X}^t \times \Theta^t$, his t -th period choice is given by

$$x_t^b(g) (x^{t-1}, \theta^t) = g_{T-(t-1)}(R^b x_{t-1}, R^b I_0 \theta_t)$$

The value of the optimization problem of agent a is then given by¹⁶

$$V_g^T(R^a x_0, R^a I_0 \theta_1) = \max_{(x_t^a)_{t=1}^T} E \left[\sum_{t=1}^T \beta^{t-1} u \left(x_{t-1}^a, x_t^a, \{x_t^b(g)\}_{b \in N(a)}, \theta_t^a \right) \right]$$

The value function associated with this dynamic choice problem can be shown to satisfy Bellman's Principle of Optimality by standard arguments (see e.g., Stokey and Lucas (1989)). It can be written in the following recursive form,

$$\begin{aligned} & V_g^{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t) \\ &= \max_{x_t^a \in X} E \left[u \left(x_{t-1}^a, x_t^a, \{x_t^b(g)\}_{b \in N(a)}, \theta_t^a \right) + \beta V_g^{T-t} \left(R^a \left(x_t^a, \{x_t^b(g)\}_{b \neq a} \right), R^a I_0 \theta_{t+1} \right) \right] \end{aligned} \quad (2)$$

for $t = 1, \dots, T$ and for all $(x^{t-1}, \theta^t) \in \mathbf{X}^t \times \Theta^t$.¹⁷ We are now ready to define our equilibrium concept.

¹⁴Heterogeneity can be incorporated into the probabilistic structure of the types θ_t^a . Also, we can allow for heterogeneity of the network structure across agents by augmenting the strategy spaces to incorporate network structure into individual heterogeneity. We explain how we do this in Section 7.1.

¹⁵That is, $c \in N(a)$ if and only if $c + (b - a) \in N(b)$. Of course, we let $\mathbb{A}$ be a linear space when we study symmetric interactions, typically $\mathbb{A} := \mathbb{Z}^d$ the d -dimensional integer lattice.

¹⁶The preference shocks being serially uncorrelated, we do not need to condition on the value of past realizations. See Section 7.2 for a treatment of persistent shocks.

¹⁷We have adopted the the convention that $V_g^0(x, I_0 \theta) := 0$ for any $(x, \theta) \in \mathbf{X} \times \Theta$.

Definition 1 A symmetric Markov Perfect Equilibrium (MPE) of a dynamic economy with social interactions is a measurable map $g^* : \mathbf{X} \times \Theta^{I(0)} \times \{1, \dots, T\} \rightarrow X$ such that for all $a \in \mathbb{A}$, for all $t = 1, \dots, T$, and for all $(x^{t-1}, \theta^t) \in \mathbf{X}^t \times \Theta^t$

$$g_{T-(t-1)}^* (R^a x_{t-1}, R^a I_0 \theta_t) \in \arg \max_{x_t^a \in X} E \left[u \left(x_{t-1}^a, x_t^a, \left\{ x_t^b(g^*) \right\}_{b \in N(a)}, \theta_t^a \right) + \beta V_{g^*}^{T-t} \left(R^a \left(x_t^a, \left\{ x_t^b(g^*) \right\}_{b \neq a} \right), R^a I_0 \theta_{t+1} \right) \right] \quad (3)$$

Clearly, an MPE is necessarily a *subgame perfect equilibrium*; that is, each agent's continuation strategy is a best response to other agent's continuation strategies after any possible history. Notice also the time notation we use for the Markovian policy: $g_{T-(t-1)}^*$ denotes the first-period equilibrium choice in a $T-(t-1)$ -periods economy. Since economies are nested, $g_{T-(t-1)}^*$ represents also the t -period equilibrium choice in a T -periods economy.

We conclude this section with a few remarks to justify our focus on MPEs. First of all, Markovian strategies are not a restriction for finite-horizon economies: we prove that the unique symmetric subgame perfect equilibrium for any finite-horizon economy is necessarily Markovian. Moreover, in an infinite horizon economy ($T = \infty$), a symmetric MPE is not necessarily stationary. The sequence of unique MPEs for finite horizon economies converges however to a $g^* : \mathbf{X} \times \Theta^{I(0)} \rightarrow X$ which turns out to be a stationary MPE of the infinite-horizon economy whose properties we focus on. Finally, we refer to Bisin, Horst and Özgür (2006) for a discussion of non-Markovian equilibria in a related context.

3 Dynamic Linear Economies with Social interactions and Conformity Preferences

We focus our attention on *linear economies* with *conformity preferences*. These are environments in which each agent's preferences incorporate the desire to conform to the choices of agents in his reference group.¹⁸

Preferences for conformity arguably provide a rationale for several important social phenomena. The empirical literature has for instance documented preferences for conformity as a motivation for smoking and other risky behaviour in teens. Similarly, the role of conformity is also documented by Glaeser, Sacerdote, and Scheinkman (1996) with regards to criminal activity and by a large literature with regards to peer effects in education outcomes.¹⁹ Conformity also represents a natural environment in which to study dynamic equilibrium. In many relevant social

¹⁸While we model preferences for conformity directly as a preference externality, we intend this as a reduced form of models of behavior in groups which induce indirect preferences for conformity, as e.g., Jones (1984), Cole, Mailath and Postlewaite (1992), Bernheim (1994), Peski (2007).

¹⁹See the Introduction for the relevant references.

phenomena, in fact, the effects of preferences for conformity are amplified by the presence of limits to the reversibility of dynamic choices. This is of course the case for smoking, alcohol abuse and other risky teen behaviour, which are hard to reverse because they might lead to chemical addictions. In other instances, while addiction per se is not at issue, nonetheless behavioural choices are hardly freely reversible because of various social and economic constraints, as is the case, for instance, of engaging in criminal activity. Finally, exogenous and predictable changes in the composition of groups, as e.g., in the case of school peers at the end of a school cycle, introduce important non-stationarities in the agents' choice. These non-stationarities also call for a formal analysis of dynamic social interactions.

With the objective of providing a clean and simple analysis of dynamic social interactions in a conformity economy, we impose strong(er than required) but natural assumptions.²⁰ In particular (i) we restrict the neighborhood correspondence to represent the minimal interaction structure allowing for overlapping groups, (ii) we restrict preferences to be quadratic, and (iii) we impose enough regularity conditions on the agents' choice problem to render it convex. Formally,

Assumption 1 *A linear conformity economy satisfies the following.*

1. Let $\mathbb{A} := \mathbb{Z}$ represent a general social space. Each agent interacts with his immediate neighbors, i.e., for all $a \in \mathbb{A}$, $N(a) := \{a - 1, a + 1\}$.
2. The contemporaneous preferences of an agent $a \in \mathbb{A}$ are represented by the utility function

$$u(x_{t-1}^a, x_t^a, x_t^{a-1}, x_t^{a+1}, \theta_t^a) := -\alpha_1(x_{t-1}^a - x_t^a)^2 - \alpha_2(\theta_t^a - x_t^a)^2 - \alpha_3(x_t^{a-1} - x_t^a)^2 - \alpha_3(x_t^{a+1} - x_t^a)^2 \quad (4)$$

where α_1, α_2 , and α_3 , are positive constants.

3. Let $X = \Theta = [\underline{x}, \bar{x}] \subset \mathbb{R}$, where $\underline{x} < \bar{x}$. Let v be absolutely continuous with a positive density²¹, $E[\theta_t^a] = \int \theta_t^a d\nu =: \bar{\theta} \in (\underline{x}, \bar{x})$, and $\text{Var}(\theta_t^a) = \int (\theta_t^a - \bar{\theta})^2 d\nu < \infty$.²²

Assumption 1-1 requires that the reference group of each agent $a \in \mathbb{A}$ be composed of his immediate neighbors in the social space, namely the agents $a - 1$ and $a + 1$. The utility function u defined in Assumption 1-2 describes the trade-off that agent $a \in \mathbb{A}$ faces between matching his individual characteristics (x_{t-1}^a, θ_t^a) and the utility he receives from conforming to the current choices of his peers (x_t^{a-1}, x_t^{a+1}) . The different values of α_i represent different levels of intensity of the social interaction motive relative to the own (or intrinsic) motive. Finally, Assumption 1-2 and 1-3 jointly guarantee that the agents' choice problem is convex.

²⁰See Section 7 for possible directions in which the structure and the results we obtain are easily generalized.

²¹We will call a measure μ 'absolutely continuous' if it is absolutely continuous with respect to the Lebesgue measure λ , i.e., if $\mu(A) = 0$ for every measurable set A for which $\lambda(A) = 0$.

²²We need absolute continuity only when we prove inefficiency. All other results are obtained without that assumption.

3.1 Equilibrium

We provide here the basic theoretical results regarding our dynamic linear social interaction economy with conformity. The reader only interested in the characterization can skip this section, keeping in mind that equilibria exist (for finite economies they are unique) and they induce an ergodic stochastic process over paths of action profiles. Furthermore, a stationary ergodic distribution also exists for the economy. Finally, a recursive algorithm to compute equilibria is derived. The proofs of all statements can be found in the Appendices.

Theorem 1 (Existence - Complete Information) *Consider an economy with conformity preferences and complete information.*

1. *If the time horizon is finite ($T < \infty$), then the economy admits a unique symmetric MPE $g^* : \mathbf{X} \times \Theta \times \{1, \dots, T\} \mapsto X$ such that for all $t \in \{1, \dots, T\}$, for all $(x_{t-1}, \theta_t) \in \mathbf{X} \times \Theta$*

$$g_{T-(t-1)}^*(x_{t-1}, \theta_t) = \sum_{a \in \mathbb{A}} c_{T-(t-1)}^a x_{t-1}^a + \sum_{a \in \mathbb{A}} d_{T-(t-1)}^a \theta_t^a + e_{T-(t-1)} \bar{\theta} \quad \mathbb{P} - a.s.$$

where $c_\tau^a, d_\tau^a, e_\tau \geq 0$, $a \in \mathbb{A}$, and $e_\tau + \sum_{a \in \mathbb{A}} (c_\tau^a + d_\tau^a) = 1$, $0 \leq \tau \leq T$. Moreover, the equilibrium is also unique in the class of subgame perfect equilibria (SPE), meaning that there does not exist any non-Markovian SPE for our economy.

2. *If the time horizon is infinite ($T = \infty$), then the economy admits a symmetric stationary MPE $g^* : \mathbf{X} \times \Theta \mapsto X$ such that*

$$g^*(x_{t-1}, \theta_t) = \sum_{a \in \mathbb{A}} c^a x_{t-1}^a + \sum_{a \in \mathbb{A}} d^a \theta_t^a + e \bar{\theta} \quad \mathbb{P} - a.s.$$

where $c^a, d^a, e \geq 0$, for $a \in \mathbb{A}$, and $e + \sum_{a \in \mathbb{A}} (c^a + d^a) = 1$.²³

The theorems in this section can be extended with straightforward modifications to the case of incomplete information. We state without proof, e.g., the existence theorem for economies with incomplete information next.

Theorem 2 (Existence - Incomplete Information) *Consider an economy with conformity preferences and with incomplete information.*

1. *For $T < \infty$, the economy admits a unique symmetric MPE $g^* : \mathbf{X} \times \Theta^{I(0)} \times \{1, \dots, T\} \mapsto X$ such that for all $t \in \{1, \dots, T\}$,*

$$g_{T-(t-1)}^*(x_{t-1}, I_0 \theta_t) = \sum_{a \in \mathbb{A}} c_{T-(t-1)}^a x_{t-1}^a + \sum_{a \in I(0)} d_{T-(t-1)}^a \theta_t^a + e_{T-(t-1)} \bar{\theta} \quad \mathbb{P} - a.s.$$

where $c_\tau^a, d_\tau^a, e_\tau \geq 0$ and $e_\tau + \sum_{a \in \mathbb{A}} c_\tau^a + \sum_{a \in I(0)} d_\tau^a = 1$, $0 \leq \tau \leq T$.

²³Several assumptions can be relaxed while guaranteeing existence. In particular, the symmetry of the neighborhood structure can be substantially relaxed. See Section 7.1 for the discussion.

2. For $T = \infty$, the economy admits a symmetric MPE $g^* : \mathbf{X} \times \Theta^{I(0)} \mapsto X$ such that

$$g^*(x_{t-1}, I_0 \theta_t) = \sum_{a \in \mathbb{A}} c^a x_{t-1}^a + \sum_{a \in I(0)} d^a \theta_t^a + e \bar{\theta} \quad \mathbb{P} - a.s.$$

where $c^a, d^a, e \geq 0$ and $e + \sum_{a \in \mathbb{A}} c^a + \sum_{a \in I(0)} d^a = 1$.

3.1.1 A Sketch of the Proof

The proof of the existence theorem requires some subtle arguments. While referring to the Appendix for details, a few comments here in this respect will be useful. Consider the (infinite dimensional) choice problem of each agent $a \in \mathbb{A}$. To be able to apply standard variational arguments to this problem it is necessary to bound the marginal effect of any infinitesimal change dx^a on the agent's value function. To this end, the Envelope theorem (as e.g., in Benveniste and Scheinkman (1979)) is not enough, as dx^a affects agent a 's value function directly and indirectly, through its effects on all agents $b \in \mathbb{A} \setminus a$'s choices, which in turn affect agent a 's value function. The marginal effect of any infinitesimal change dx^a is then an infinite sum. Furthermore, each term in the sum contains endogenous terms from some agent $b \in \mathbb{A} \setminus a$'s policy function (and there is an infinite number of them), which makes it impossible to adopt the methodology used by Santos (1991) to prove the smoothness of the policy function in infinite dimensional recursive choice problems. In our economy, with quadratic utility, policy functions are necessarily linear and, provided we show that equilibria are interior, symmetric MPE's can be represented by a policy function which is obtained as a fixed point of a recursive map which can be directly studied. Extending the existence proof to general preferences would require therefore sufficient conditions on the structural parameters to control the curvature of the policy function of each agent's decision problem. We conjecture that this can be done although sufficient conditions do not appear transparently from our proof. A more detailed sketch of the steps involved in the existence proof follows.

Step 1 In the last period (1-period continuation) of any finite-horizon economy, first order conditions (FOC's) induce a contraction operator on the space of bounded measurable functions having as arguments any t -length history. Hence, there exists a unique symmetric (possibly history-dependent) equilibrium. We then show that the equilibrium policy must be Markovian and should take the convex combination form in the statement of Theorem 1 (Lemma 1).

Step 2 For any finite horizon ($T < \infty$) economy, we assume that in the continuation from period 2 on agents choose according to the unique symmetric MPE, $g : \mathbf{X} \times \Theta \times \{1, \dots, T-1\} \mapsto X$. Linearity of the policy in the continuation keeps a generic agent's dynamic program strictly concave and FOC's are necessary and sufficient for a pure strategy maximum. We show

that we can write FOC's as functions only of first period choices and preference and the mean preference shocks. By the same token as in Step 1, we focus on Markovian strategies. FOC induces a contraction operator on the set of Markovian strategies into itself. Hence, there exists a unique fixed point $g_T^* : \mathbf{X} \times \Theta \mapsto X$. We conclude that for the T -period economy, the map $(g_T^*, g) : \mathbf{X} \times \Theta \times \{1, 2, \dots, T\} \mapsto X$ is the unique symmetric MPE in pure strategies and has the convex combination form as in the statement of the theorem, which completes the induction argument.

Step 3 The final step involves taking a limit. We construct a series of finite economies, approximating the ∞ -horizon economy, given an appropriate topology. We then show that, the finite truncation equilibrium correspondence is upper-hemi-continuous (u.h.c.) with respect to the parametrization. This is however not enough for stationarity. We prove that the behavioral Markovian strategy set (the set G) is compact. This helps us prove that the sequence of finite-horizon equilibrium policy functions converges uniformly to a policy function in G (which is an equilibrium policy due to u.h.c of the equilibrium correspondence), hence the same one every period, after any history. This gives us stationarity.

3.2 The parameters of the policy function

By exploiting the linearity of policy functions, our method of proof is constructive, producing a direct and useful recursive computational characterization for the parameters of the symmetric policy function at equilibrium. We repeatedly exploit this characterization in the next section e.g., when performing comparative dynamics exercises. Consider the choice problem of agent 0. For any T -period economy, agent 0's dynamic program yields a FOC that takes the following form (see Lemma 3)

$$x_1^0 = \Delta_T^{-1} \left(\alpha_1 x_0^0 + \alpha_2 \theta_1^0 + \sum_{a \neq 0} \gamma_T^b x_1^b + \mu_T \bar{\theta} \right) \quad (5)$$

where Δ_T and γ_T^b , and μ_T are the effects on agent zero's discounted expected marginal utility of changes in agents 0 and b 's first period actions and the change in the level of $\bar{\theta}$, respectively. Let

$$L_{c,d,e} := \{(c, d, e) : e \geq 0, c^a \geq 0, d^a \geq 0, \forall a \text{ and } e + \sum_a (c^a + d^a) = 1\}$$

be the space of nonnegative coefficient sequences whose sum is 1. The existence of an equilibrium policy for the first period of a T -period economy is then equivalent to the existence of a coefficient sequence (c_T^*, d_T^*, e_T^*) which is the fixed point of a map $L_T : L_{c,d,e} \rightarrow L_{c,d,e}$ induced by (5) s.t.

$(\hat{c}, \hat{d}, \hat{e}) = L_T(c, d, e)$ and for each $a \in \mathbb{A}$

$$\begin{aligned}\hat{c}^a &= \Delta_T^{-1} \left(\alpha_1 \mathbf{1}_{\{a=0\}} + \sum_{b \neq 0} \gamma_T^b c^{a-b} \right) \\ \hat{d}^a &= \Delta_T^{-1} \left(\alpha_2 \mathbf{1}_{\{a=0\}} + \sum_{b \neq 0} \gamma_T^b d^{a-b} \right) \\ \hat{e} &= \Delta_T^{-1} \left(\mu_T + e \sum_{b \neq 0} \gamma_T^b \right)\end{aligned}\tag{6}$$

by matching coefficients of the policy on both sides of (5). The parameters of the map L_T , namely $\Delta_T, (\gamma_T^b)_{a \neq 0}, \mu_T$, depend only on the continuation equilibrium coefficients $(c_s^*, d_s^*, e_s^*)_{s=1}^{T-1}$ in a linear fashion (see (31), (48), and (50) for their detailed expressions). For $T = 1$, the parameters of L_1 are dictated directly by the underlying preferences, namely $\Delta_1 = \alpha_1 + \alpha_2 + 2\alpha_3$, $\gamma_1^1 = \gamma_1^{-1} = \alpha_3$, $\gamma_1^b = 0$, for all $b \neq -1, 0, 1$, and $\mu_1 = 0$. Thus, the map L_1 defined by the system in (6) becomes

$$\begin{aligned}\hat{c}^a &= \Delta_1^{-1} (\alpha_1 \mathbf{1}_{\{a=0\}} + \alpha_3 c^{a-1} + \alpha_3 c^{a+1}) \\ \hat{d}^a &= \Delta_1^{-1} (\alpha_2 \mathbf{1}_{\{a=0\}} + \alpha_3 d^{a-1} + \alpha_3 d^{a+1}) \\ \hat{e} &= \Delta_1^{-1} (2\alpha_3 e)\end{aligned}\tag{7}$$

which is a contraction mapping whose unique fixed point is computed as the unique root to a second-order difference equation that satisfies transversality conditions toward both infinities. Consequently, the equilibrium policy coefficients are computed as in the next Theorem.

Theorem 3 (Recursive algorithm) *Consider a finite-horizon T -period economy with conformity preferences $(\alpha_i > 0, i = 1, 2, 3)$ and complete information.*

- (i) *The map L_1 for a one-period economy, defined in (6), forms a second-order difference equation for the equilibrium coefficient sequence, whose unique non-explosive, exponential solution is the unique fixed point of L_1 . We compute the coefficient sequence in closed-form. For any $a \in \mathbb{A}$,*

$$c_1^{*a} = r_1^{|a|} \left(\frac{\alpha_1}{\alpha_1 + \alpha_2} \right) \left(\frac{1 - r_1}{1 + r_1} \right) \quad \text{and} \quad d_1^{*a} = r_1^{|a|} \left(\frac{\alpha_2}{\alpha_1 + \alpha_2} \right) \left(\frac{1 - r_1}{1 + r_1} \right)\tag{8}$$

where

$$r_1 = \left(\frac{\Delta_1}{2\alpha_3} \right) - \sqrt{\left(\frac{\Delta_1}{2\alpha_3} \right)^2 - 1}$$

with $\Delta_1 = \alpha_1 + \alpha_2 + 2\alpha_3$.

- (ii) *The coefficients $(c_s^*, d_s^*, e_s^*)_{s=2}^T$ of the sequence of Markov equilibrium policies are computed recursively as the unique fixed points of the recursive contraction maps $L_s : L_{c,d,e} \rightarrow L_{c,d,e}$, $s = 2, \dots, T$, defined in (6), whose parameters $\Delta_s, (\gamma_s^a)_{a \neq 0}, \mu_s$ depend linearly only on the continuation equilibrium policy coefficients $(c_\tau^*, d_\tau^*, e_\tau^*)_{\tau=1}^{s-1}$, as defined in (31), (48), and (50).*

(iii) Moreover, $\lim_{T \rightarrow \infty} (c_T^*, d_T^*, e_T^*) = (c^*, d^*, e^*)$ exists and it is the coefficient sequence of the stationary Markovian equilibrium policy function for the infinite-horizon economy whose existence is proved in Theorem 1.

Fixed point calculations take less than a few seconds on an ordinary computer, for each period. Finally, the sequence of fixed point maps that we compute at each iteration converges to a policy sequence, which turns out to be the infinite-horizon stationary MPE. The convergence is very rapid, under a few minutes.

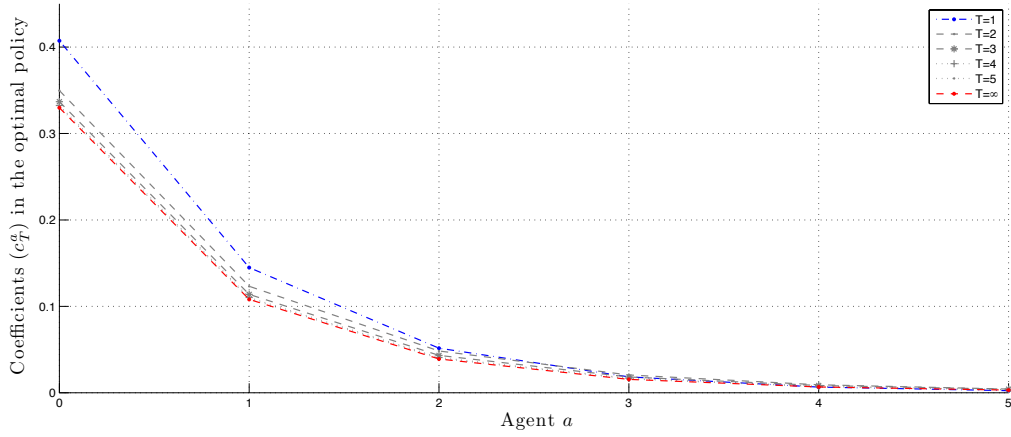


Figure 1: Non-stationary Optimal Policy.

3.3 Ergodicity

With such characterization of the parameters of the policy function at hand, we are able to characterize very tightly the spatial (cross-sectional) and intertemporal behavior of the equilibrium process emerging from the class of dynamic models we study. Let π_0 be an initial distribution on the configuration space $\mathbf{X}$. Given the initial distribution π_0 , a stationary MPE of the economy with conformity induces an equilibrium process $(x_t \in \mathbf{X})_{t=0}^\infty$ (via the policy function g^*) and an associated transition function Q_{g^*} . This latter generates iteratively a sequence of distributions $(\pi_t)_{t=1}^\infty$ on the configuration space $\mathbf{X}$, i.e., for $t = 0, 1, \dots$

$$\pi_{t+1}(A) = \pi_t Q_{g^*}(A) = \int_{\mathbf{X}} Q_{g^*}(x_t, A) \pi_t(dx_{t+1}) \quad (9)$$

We show first that, given the induced equilibrium process, the transition function Q_{g^*} admits an *invariant distribution* π , i.e., $\pi = \pi Q_{g^*}$ and that the stationary equilibrium process starting from

π is *ergodic*²⁴.

Ergodicity does not necessarily imply the convergence of the equilibrium process to a unique distribution starting from an arbitrary initial distribution π_0 . Conditions are necessary to guarantee such convergence.²⁵ We next show that, for any initial distribution π_0 and a stationary Markovian policy function g^* , the equilibrium process $(x_t \in \mathbf{X})_{t=0}^\infty$ converges in distribution to the invariant distribution π , independently of π_0 .²⁶ This also implies that π is the *unique* invariant distribution of the equilibrium process $(x_t \in \mathbf{X})_{t=0}^\infty$. More specifically,

Theorem 4 (Ergodicity) *Suppose the process $((\theta_t^a)_{t=-\infty}^\infty)_{a \in \mathbb{A}}$ is i.i.d. with respect to a and t according to ν . The equilibrium process $(x_t \in \mathbf{X})_{t=0}^\infty$ induced by a symmetric stationary Markov perfect equilibrium of an economy with conformity via the policy function $g^*(x_{t-1}, \theta_t)$ and the unique invariant measure π as the initial distribution is ergodic; π is the joint distribution of*

$$x_t = \left(\frac{e^{\bar{\theta}}}{1-C} + \sum_{s=1}^{\infty} \sum_{b_1 \in \mathbb{A}} \cdots \sum_{b_s \in \mathbb{A}} c^{b_1} \cdots c^{b_{s-1}} d^{b_s} \theta_{t+1-s}^{a+b_1+\cdots+b_s} \right)_{a \in \mathbb{A}} \quad (10)$$

where $C := \sum_{a \in \mathbb{A}} c^a$ is the sum of coefficients in the stationary policy function that multiply corresponding agents' last period choices. Moreover, the sequence $(\pi_t)_{t=1}^\infty$ of distributions generated by the equilibrium process $(x_t \in \mathbf{X})_{t=0}^\infty$ converges to π in the topology of weak convergence for probability measures, independently of any arbitrary initial distribution π_0 .²⁷

4 Characterization of equilibrium

Exploiting the linear structure of our economies we can study equilibria in some detail. Recall that the policy function in each period $t = 1, \dots, T$, for each agent $a \in \mathbb{A}$, is

$$x_t^a = \sum_{b \in \mathbb{A}} c_{T-(t-1)}^b x_{t-1}^{a+b} + \sum_{b \in \mathbb{A}} d_{T-(t-1)}^b \theta_t^{a+b} + e_{T-(t-1)} \bar{\theta}, \quad (11)$$

with $e_{T-(t-1)} + \sum_{a \in \mathbb{A}} (c_{T-(t-1)}^a + d_{T-(t-1)}^a) = 1$, when T is finite; and

²⁴We call a Markov process (x_t) with state space $\mathbf{X}$ under a probability measure P ergodic if $\frac{1}{T} \sum_{t=1}^T f(x_t) \rightarrow \int f dP$ P -almost surely for every bounded measurable function $f : \mathbf{X} \rightarrow \mathbb{R}$. See for example Blume (1982), Duffie et al (1994) and Hansen (1982) for the use of ergodicity in dynamic economic theory and modern econometric theory.

²⁵The well-known Döblin conditions to that effect can be found in Doob (1953). See also Futia (1982), Neveu (1965), and Tweedie (1975) for similar characterizations.

²⁶Note however that Theorem 1 does not guarantee that the policy function $g^*(x_{t-1}, \theta_t)$ is unique.

²⁷A sequence of probability measures (λ_t) is said to converge weakly (or in the topology of weak convergence for probability measures) to λ if, for any bounded, measurable, continuous function $f : \mathbf{X} \rightarrow \mathbb{R}$, $\lim_{t \rightarrow \infty} \int f d\lambda_t = \int f d\lambda$ almost surely (see e.g. Kallenberg (2002), p.65).

$$x_t^a = \sum_{b \in \mathbb{A}} c^b x_{t-1}^{a+b} + \sum_{b \in \mathbb{A}} d^b \theta_t^{a+b} + e \bar{\theta}, \quad (12)$$

with $e + \sum_{a \in \mathbb{A}} (c^a + d^a) = 1$, in the infinite-horizon case.

First of all, we study the *parameters of the policy function*. The coefficients $c_{T-(t-1)}^b$ and $d_{T-(t-1)}^b$ (resp. c^b and d^b in the case of infinite-horizon economies), in particular, may be viewed as a measure for the total impact of the action x_{t-1}^{a+b} and of the preference shock θ_t^{a+b} of agent $a+b$, respectively, on the optimal current choice of agent a ; where b concisely represents the *social distance* between the two agents.²⁸ Furthermore, we study a fundamental statistical property of equilibrium, *cross-sectional auto-correlation of actions*. In fact, although any agent $a \in \mathbb{A}$ interacts directly only with a small subset of the population, at equilibrium, each agent's optimal choice is correlated with those of all the other agents. Let $\rho_{a,T}$ denote the conditional correlation between the first-period equilibrium actions of agents a -step away from each other, in the T -period economy, given $x_0 \in \mathbf{X}$:²⁹

$$\rho_{a,T} = \frac{\text{Cov}(x_1^0, x_1^a \mid x_0)}{\text{Var}(x_1^0 \mid x_0)}. \quad (13)$$

4.1 Policy Function

Consider first a finite-horizon economy. Since the policy function for this economy is well-defined, the coefficients $c_{T-(t-1)}^b$ and $d_{T-(t-1)}^b$ satisfy

$$\lim_{|b| \rightarrow \infty} c_{T-(t-1)}^{a+b} = \lim_{|b| \rightarrow \infty} d_{T-(t-1)}^{a+b} = 0$$

The impact of an agent $a+b$ on agent a tends to zero as $|b| \rightarrow \infty$. In this sense, linear conformity economies display weak social interactions.

Furthermore, as we have shown in Section 3.2,

$$\lim_{T \rightarrow \infty} c_T = c, \quad \lim_{T \rightarrow \infty} d_T = d, \quad \text{and} \quad \lim_{T \rightarrow \infty} e_T = e$$

²⁸See Akerlof (1997) for richer definitions of social distance.

²⁹The correlation between the first-period optimal choices of agents a and b , is

$$\frac{\text{Cov}(x_1^a, x_1^b \mid x_0)}{\sqrt{\text{Var}(x_1^a \mid x_0) \text{Var}(x_1^b \mid x_0)}}.$$

Due to the symmetry imposed on our economy, such correlations are independent of agents' labels but depends only on $|b - a|$. Consequently, we can define the conditional correlation function with distances computed relative to any agent, in particular agent 0.

The finite-horizon parameters converge (uniformly) to the infinite-horizon stationary policy parameters.

Finally, equilibrium policy functions are non-stationary in the finite economy, as rational forward-looking agents change their behavior optimally through time. In the final periods, for example, social interactions lose weight relative to individual characteristics; see Figure 1.³⁰

4.2 Cross-sectional Auto-correlations

Exploiting the equilibrium characterization provided by Theorems 1 and 3, and the independence of preference shocks across agents, we can compute the covariance terms:

$$Cov\left(x_1^0, x_1^a \middle| x_0\right) = Var(\theta) \sum_{a_1 \in \mathbb{A}} d_T^{a_1} d_T^{a_1-a}. \quad (14)$$

The expression $\sum_{a_1 \in \mathbb{A}} d_T^{a_1} d_T^{a_1-a}$ is the discrete self-convolution of the equilibrium policy sequence $d_T = (d_T^{a_1})_{a_1 \in \mathbb{A}}$, where a acts as the shift parameter.³¹ In Figure 2 we show how the convolution behaves with respect to the distance a , for the same set of parameters as in Figure 1. Substituting the form in (14) back in (13) for both terms, we obtain

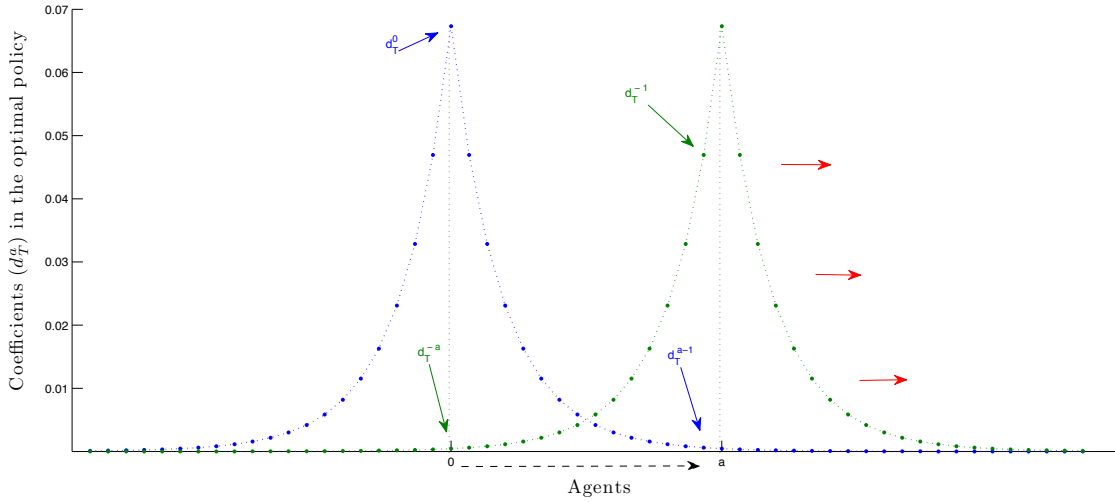


Figure 2: Convolution of the Policy Coefficient Sequence.

³⁰We plot in Figure 1 only one side of the policy coefficient sequence to get a close-up view of the change in equilibrium behavior. The left hand side is the mirror image of that due to symmetry. Parameter values for this figure are $\frac{\alpha_1}{\alpha_3} = 1$, $\frac{\alpha_3}{\alpha_2} = 10$, and $\beta = .95$

³¹See (53) for the derivation.

$$\rho_{a,T} = \frac{\sum_{b \in \mathbb{A}} d_T^b d_T^{b-a}}{\sum_{b \in \mathbb{A}} d_T^b d_T^b} \quad (15)$$

the a -step conditional cross-sectional autocorrelations for the first-period equilibrium choices of the T -period economy. Exploiting the recursive algorithm provided by Theorem 3, we can compute these autocorrelations easily for any finite economy. We can then study the behavior of the conditional correlation function $\rho_{a,T}$ through time (T) and across social space (a). These correlations exhibit interesting dynamics: they are declining in a , for any T , but the rate of decline cannot be ranked in T , given a ; see Figure 3 for an example with the same parametrization we used above for the policy weights in Figure 1. In particular, given a T -period economy, consider

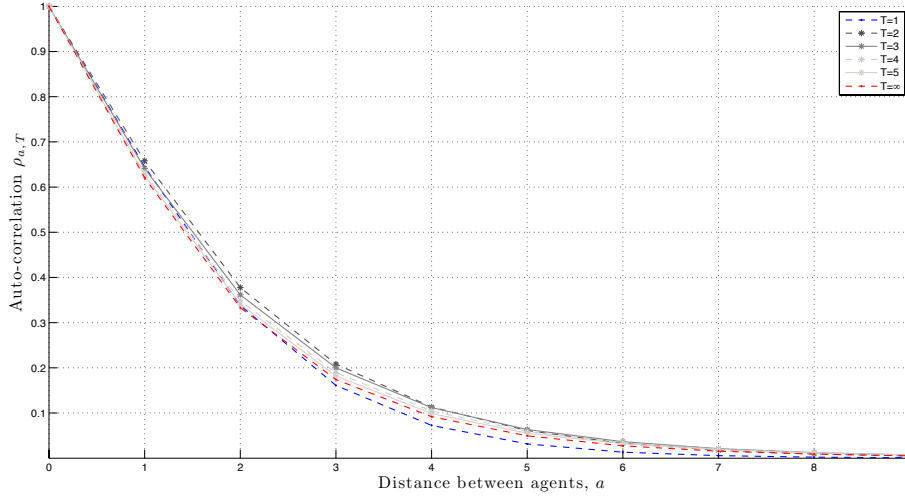


Figure 3: Cross-sectional Auto-correlations.

the T -period rate of convergence of the spatial autocorrelations, for $a \geq 0$ ³²

$$r_{a,T} = \frac{\rho_{a+1,T}}{\rho_{a,T}}.$$

It is easy to show analytically that $r_{a,1}$ declines monotonically and becomes constant at the tail in a .³³ On the other hand, $r_{a,T}$ is typically non-monotonic in a , for longer horizons, including for $T = \infty$; see Figure 4.

Finally, consider the T -period rate of tail convergence of the spatial autocorrelations,

$$r_T := \lim_{a \rightarrow \infty} r_{a,T} = \lim_{a \rightarrow \infty} \left(\frac{\rho_{a+1,T}}{\rho_{a,T}} \right)$$

Similarly, let the same rate for the infinite-horizon economy ($T = \infty$) be represented by r .

³²The rate is symmetrically defined with respect to agent 0, i.e., $r_{a,T} = \frac{\rho_{a-1,T}}{\rho_{a,T}}$, for any $a \leq 0$.

³³See the proof of Proposition 1 for the argument.

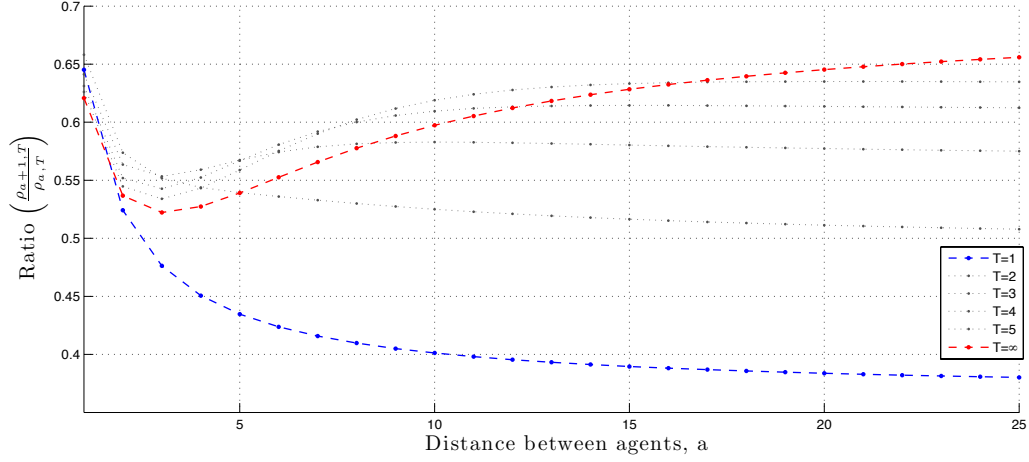


Figure 4: Rate of Convergence of the Auto-correlations.

Proposition 1 (Tail Convergence Monotonicity) ³⁴ *The rate r_T is monotone increasing with respect to the length of the economy,*

$$r_{T+1} > r_T, \text{ for finite } T \geq 1.$$

Moreover, the sequence of tail convergence rate for finite-horizon economies converges to that of the infinite-horizon economy as the horizon length gets larger and the limit rate is strictly less than 1:

$$\lim_{T \rightarrow \infty} r_T = r < 1.$$

In other words, even though the autocorrelation functions might behave non-monotonically for shorter social distances, they eventually converge (as social distance $a \rightarrow \infty$) to an exponential rate in the tail. Moreover, rates of tail convergence are higher the farther is the final period of the economy (as $T \rightarrow \infty$). This is because rational agents choose to correlate their actions more with their neighbors in early periods and progressively less so as they approach the end of their social interactions. Finally, as the infinite-horizon limit is approached, the rate of tail convergence becomes stationary (as to be expected since finite-horizon equilibria approximate the stationary infinite-horizon equilibrium). We use this intuition to the fullest extent when discussing identification in Section 6.

In an infinite-horizon economy social interactions manifest themselves at the stationary ergodic distribution by means of spatial autocorrelation of actions. Given $x_0 \in \mathbf{X}$, the conditional covariance in period t of an infinite-horizon economy, between two agents a agents away from each

³⁴The proof is in Appendix D.

other is denoted by $Cov(x_t^0, x_t^a \mid x_0)$. Let $Cov(x^0, x^a)$ be the a -step *unconditional* covariance at the ergodic stationary distribution. Since the stationary MPE is ergodic, it is easy to see from Lemma 2 (i) and Theorem 4 that as t gets arbitrarily large, the conditional t -period covariance between agents 0 and a converges to its unconditional counterpart at the limit distribution, i.e.,

$$Cov(x^0, x^a) = \lim_{t \rightarrow \infty} Cov(x_t^0, x_t^a \mid x_0)$$

Moreover, the limit unconditional correlation ρ_b between the actions of agents a and $a + b$ is independent of x_0 and it satisfies

$$\rho_b = \frac{Cov(x^0, x^a)}{Var(x^0)} = \lim_{t \rightarrow \infty} \frac{Cov(x_t^0, x_t^a \mid x_0)}{Var(x_t^0 \mid x_0)}$$

Finally, because of the stationarity of the policy function in (12), the limit covariance between two agents a agents away from each other can be written as

$$\begin{aligned} Cov(x^0, x^a) &= \lim_{t \rightarrow \infty} Cov(x_t^0, x_t^a \mid x_0) \\ &= \sum_{a_1 \in \mathbb{A}} \sum_{b_1 \in \mathbb{A}} c^{a_1} c^{b_1} Cov(x^{a_1}, x^{a+b_1}) + Var(\theta) \sum_{a_1 \in \mathbb{A}} d^{a_1} d^{a_1-a}, \end{aligned} \quad (16)$$

and hence it has a simple recursive structure. In fact, since the sum of the stationary weights multiplying covariances on the right hand side are strictly less than one, this system can be seen as a contraction operator. Hence, for each one-step conditional autocorrelation sequence, there is a unique stationary unconditional autocorrelation sequence that we can compute using the above recursive system easily. We later exploit this recursive structure further in Section 6.1 when we compare equilibrium stationary distributions induced by myopic and rational agents.

In Figure 5, we report the correlation functions in both the mild and strong conformity parameterizations as a function of social distance, b .³⁵ Two effects are worth mentioning here. Firstly, both correlation functions converge to zero as the distance between two agents become arbitrarily large. Secondly, this convergence is much faster in the case of mild interactions than in the case of strong interactions. For example, the correlation between the equilibrium choices of agent a and agent $a + 3$ (or $a - 3$ due to symmetry) is about 7% in the case of mild interactions whereas it is about 75% in the case of strong interactions. The correlation between the equilibrium choices of agent a and agent $a + 6$ are about 0% and 40% respectively. The strength of the desire to conform built in individuals' preferences determine endogenously, at equilibrium, the size of the effective neighborhood with which an individual interacts.

³⁵See Section 5.2 for the parameter values for the mild and strong interaction cases.

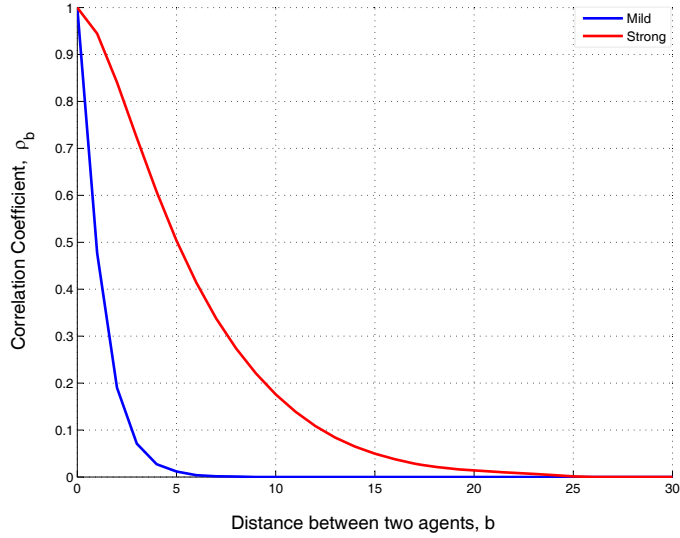


Figure 5: Correlation function at the ergodic distribution for Mild and Strong Interactions.

5 Equilibrium Properties and Comparative dynamics

In this section we first study the welfare properties of equilibrium and then we use the characterization of equilibria we obtained to produce several simulations illustrating various comparative dynamics exercises of interest.

5.1 (In)efficiency

Social interactions are modelled in this paper as a preference externality, that is, by introducing a dependence of agent a 's preferences on his/her peers' actions. Not surprisingly, therefore, equilibria will not be efficient in general. In this section we also characterize the form the inefficiency takes when social interactions are modelled as preferences for conformity.

A benevolent social planner, taking into account the preference externalities and at the same time treating all agents symmetrically, would maximize the expected discounted utility of a generic agent, say of agent $a \in \mathbb{A}$, by choosing a symmetric choice function $h \in CB(\mathbf{X} \times \Theta, X)$, the space of bounded, continuous, and X -valued measurable functions. In other words, h is the solution

to³⁶

$$\begin{aligned} \max_{\{f \in CB(\mathbf{X} \times \Theta, X)\}} \int & \left[\sum_{t=1}^T \beta^{t-1} \left(-\alpha_1 (f_{T-(t-2)}(R^a x_{t-2}, R^a I_0 \theta_{t-1}) - f_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \right. \right. \\ & -\alpha_2 (\theta_t^a - f_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \\ & -\alpha_3 (f_{T-(t-1)}(R^{a-1} x_{t-1}, R^{a-1} I_0 \theta_t) - f_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \\ & \left. \left. -\alpha_3 (f_{T-(t-1)}(R^{a+1} x_{t-1}, R^{a+1} I_0 \theta_t) - f_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \right) \right] \prod_{t=1}^T \mathbb{P}(d\theta_t) \pi_0(dx_0) \end{aligned}$$

where π_0 is an absolutely continuous distribution on the initial choice profiles with a positive density. This problem can be written recursively. For any agent $a \in \mathbb{A}$, for all $t = 1, \dots, T$, and all $(x_{T-1}, \theta_T) \in \mathbf{X} \times \Theta^{I(0)}$, let the value of using the choice rule h in the continuation be defined as

$$\begin{aligned} V^{h, T-(t-1)}(R^a x_{T-1}, R^a \theta_T) &= -\alpha_1 (x_{t-1}^0 - h_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \\ &\quad -\alpha_2 (\theta_t^a - h_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \\ &\quad -\alpha_3 (h_{T-(t-1)}(R^{a-1} x_{t-1}, R^{a-1} I_0 \theta_t) - h_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \\ &\quad -\alpha_3 (h_{T-(t-1)}(R^{a+1} x_{t-1}, R^{a+1} I_0 \theta_t) - h_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \\ &\quad + \beta \int V^{h, T-t} \left(R^a \left\{ h_t(R^b x_{t-1}, R^b I_0 \theta_t) \right\}_{b \in \mathbb{A}}, R^a I_0 \theta_{t+1} \right) \mathbb{P}(d\theta_{t+1}) \end{aligned}$$

which leads us to the following definition

Definition 2 (Recursive Planning Problem) *Let a T -period linear economy with social interactions and conformity preferences be given. Let π_0 be an absolutely continuous distribution on the initial choice profiles with a positive density. A symmetric Markovian choice function $g : \mathbf{X} \times \Theta^{I(0)} \times \{1, \dots, T\} \rightarrow X$ is said to be **efficient** if it is a solution, for all $a \in \mathbb{A}$, and for all $t = 1, \dots, T$, to*

$$\begin{aligned} \arg \max_{\{f \in CB(\mathbf{X} \times \Theta, X)\}} \int & \left[-\alpha_1 (x_{t-1}^0 - h_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \right. \\ & -\alpha_2 (\theta_t^a - h_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \\ & -\alpha_3 (h_{T-(t-1)}(R^{a-1} x_{t-1}, R^{a-1} I_0 \theta_t) - h_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \\ & -\alpha_3 (h_{T-(t-1)}(R^{a+1} x_{t-1}, R^{a+1} I_0 \theta_t) - h_{T-(t-1)}(R^a x_{t-1}, R^a I_0 \theta_t))^2 \\ & \left. + \beta V^{h, T-t} \left(R^a \left\{ h(R^b x_{t-1}, R^b I_0 \theta_t) \right\}_{b \in \mathbb{A}}, R^a I_0 \theta_{t+1} \right) \right] \mathbb{P}(d\theta_t) \mathbb{P}(d\theta_{t+1}) \pi_t(dx_{t-1}) \end{aligned}$$

where π_t is the distribution on t -th period choice profiles induced by π_0 and the planner's choice rule h .

³⁶With the convention that $f_{T-(t-2)}(R^a x_{t-2}, R^a I_0 \theta_{t-1}) = x_0^a$ when $t = 1$.

As noted, preferences for conformity introduce an externality in each agent $a \in \mathbb{A}$'s decision problem, which depends directly on the actions of agents in neighbourhood $N(a)$ and, indirectly, on the actions of all agents in the economy. In equilibrium, agents do not internalize the impact of their choices on other agents today and in the future. More precisely,

Theorem 5 (Inefficiency of equilibrium) *A symmetric MPE of a conformity economy is inefficient.*

Furthermore, an efficient policy function will tend to weight less heavily the agent's own-effect and more heavily other agents' effects, relative to the equilibrium policy. This effect, hence the inefficiency, are neatly exhibited by comparing the equations determining policy weights in the planner (36) and equilibrium (24) scenarios. The (absolute value of the) weights the planner's equation associates on neighbors' choices is twice as large as the weights associated to neighbors in the equilibrium equation ($(\frac{\alpha_3}{\alpha_1 + \alpha_2 + 2\alpha_3})$). As a consequence, the relative weights that the planner assigns to neighbors' choices are always higher than the ones that each agent uses in equilibrium.³⁷

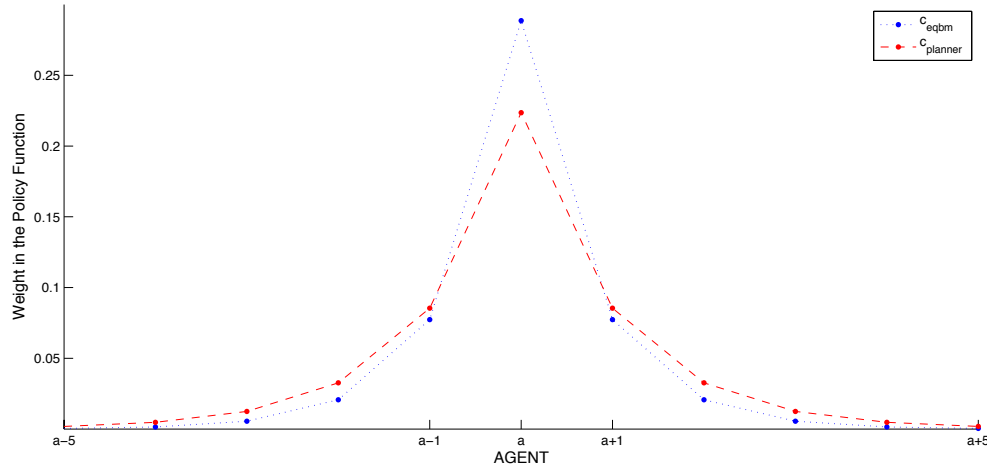


Figure 6: Inefficiency of equilibrium.

A graphic representation of the inefficiency is obtained in Figure 6, which presents the coefficient plot for the equilibrium policy of a one-period economy (equivalently the final period of any finite-horizon economy): c_{eqbm} (blue dots), and for the planner's solution, $c_{planner}$ (red dots), respectively, for a given agent $a \in \mathbb{A}$, and for a given set of parameter values ($\frac{\alpha_1}{\alpha_2} = \frac{\alpha_2}{\alpha_3} = 1$, and

³⁷Normalizing the relative coefficients to form a probability measure (see the argument in the proof of Lemma 2 (iv)), we have that the measure obtained from the planner's policy is a *mean-preserving spread* of the measure obtained at equilibrium.

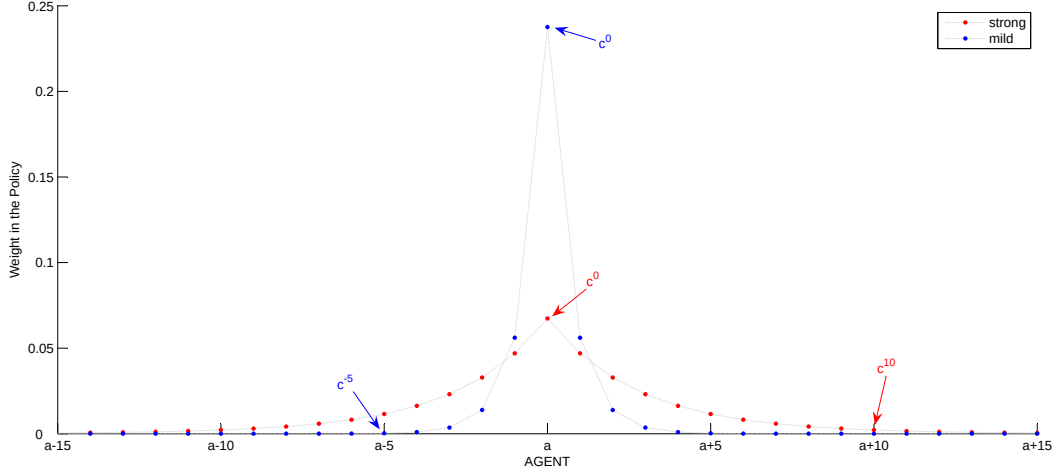


Figure 7: Weights on past history in the stationary policy function.

$\beta = .95$).³⁸

5.2 Comparative Dynamics: Peer Effects

The strength of the agents' preferences for conformity depends on the size of the preference parameter α_3 relatively to α_1 and α_2 . A policy function is represented in Figure 7, which compares a case with *mild* preferences for conformity (with parametrization $\frac{\alpha_1}{\alpha_2} = \frac{\alpha_2}{\alpha_3} = 1$)³⁹ with one with *strong* preferences for conformity (with parametrization $\frac{\alpha_1}{\alpha_2} = 1$, $\frac{\alpha_2}{\alpha_3} = \frac{1}{20}$). On the x-axis, we plot agent a and his neighbors, while on the y-axis, we plot the weights $(c^b)_{b \in \mathbb{A}}$ that the symmetric policy function g associates with the last period actions of agents $(a + b)_{b \in \mathbb{A}}$. While each agent's interaction neighborhood is only composed of two agents, in effect local interactions involve indirectly larger groups. How large are the groups depends endogenously on the strength of the agents' preferences for conformity. Notice e.g., that in Figure 7, local interactions involve effectively a group of about ten neighbors when preferences for conformity are mild and involve a group of about thirty neighbors when preferences for conformity are strong. Furthermore, for the same cases of mild and strong conformity, we compare in Figure 8 the case in which neighborhoods are overlapping, $N(a) = \{a - 1, a + 1\}$, with the case of non-overlapping one-sided neighborhoods, $N(a) = \{a + 1\}$.⁴⁰ Two effects are present here. Firstly, as in Figure 7, an increase

³⁸We call this parametrization the *mild-interaction* case in Section 5.2.

³⁹The discount rate is fixed at $\beta = .95$ in all the simulations unless mentioned otherwise.

⁴⁰In this case, the policy function is

$$x_t^a = g(R^a x_{t-1}, R^a \theta_t) = \sum_{b \geq 0} c^b x_{t-1}^{a+b} + \sum_{b \geq 0} d^b \theta_t^{a+b} + e \bar{\theta}.$$

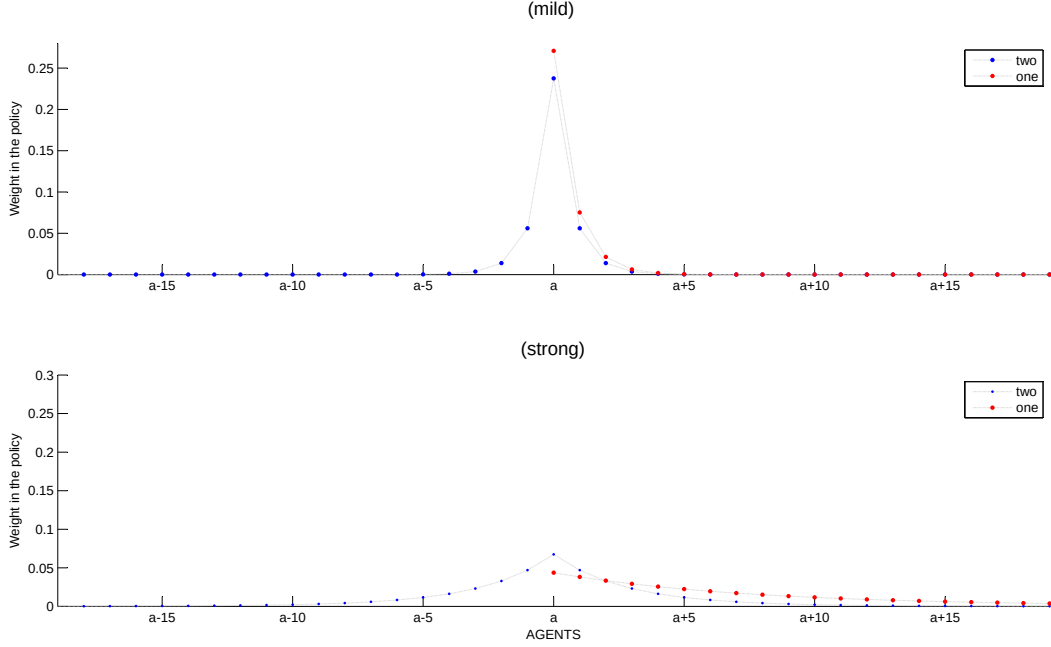


Figure 8: One-sided vs. two-sided interactions.

in the strength of the interaction parameter spreads the interaction effects over a larger social geography. Secondly, this spread is observed most significantly in the case of non-overlapping neighborhoods due to the uni-directional nature of the interactions. In turn, spatial correlations induce correlated actions of agents in endogenously formed groups.

At the ergodic stationary distribution, when the dependence of the agents' actions in equilibrium are independent of the initial configuration of actions x_0 , such correlations in endogenously formed groups is manifested in a phenomenon which we refer to as *local norms of behavior* (see Figure 9).⁴¹ In Figure 9, we plot 100 neighboring agents on the x-axis and their optimal choices drawn from the limit distribution at the same future date, on the y-axis. In the top panel, clearly the optimal actions are more spread and do not follow a significant pattern. In the bottom panel though, the optimal choices are more concentrated and follow a clear path. This is due to the fact that, in equilibrium agents conform to the actions of neighboring agents, leading the way to the creation of similar local behavior. In the bottom panel of Figure 9, we observe groups of agents (e.g., in the neighborhood of agent 20) choosing relatively low actions, and other groups (e.g., in the neighborhood of agent 70) choosing instead high actions. Two interesting aspects of this phenomenon are firstly that every individual uses the same symmetric policy function to make his choices and all heterogeneity is captured by random types and we still have high spatial

⁴¹See Appendix E for details about how we simulate the ergodic stationary distribution of actions of the economy.

correlation and high spatial variation. Secondly, the initial configuration of actions is irrelevant since the limit distribution of individual actions in this economy is ergodic.

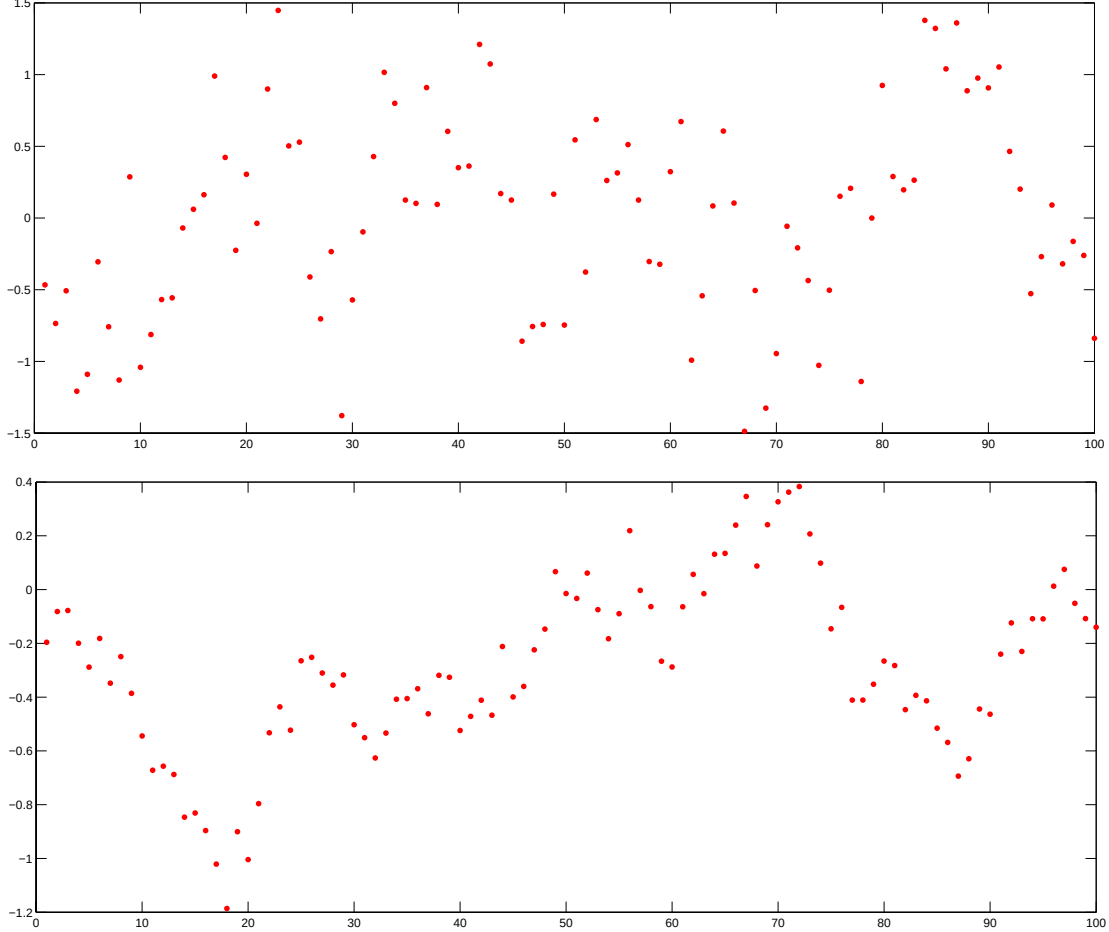


Figure 9: Ergodic Limit of *Mild* (top) and *Strong* (bottom) Interactions for 100 adjacent agents.

5.2.1 Comparative Dynamics: Information

In Figure 10, we compare the case in which agents have complete information with the case in which they have incomplete information. In this last case, the policy function is

$$x_t^a = g(R^a x_{t-1}, R^a I_0 \theta_t) = \sum_{b \in \mathbb{A}} c^b x_{t-1}^{a+b} + \sum_{b \in I(0)} d^a \theta_t^{a+b} + e \bar{\theta}.$$

In particular, we record the effect of an expansion of the information set $R^a I_0 \theta_t$ (individuals whose types are observed by agent a) on best responses. We start with an information structure in which each agent observes his own type only. We then increase the number of types observed by

each agent a (maintaining the symmetry of two-sided interactions) up to the complete information limit. The red dots represent the optimal weights in the policy function of an agent a as a

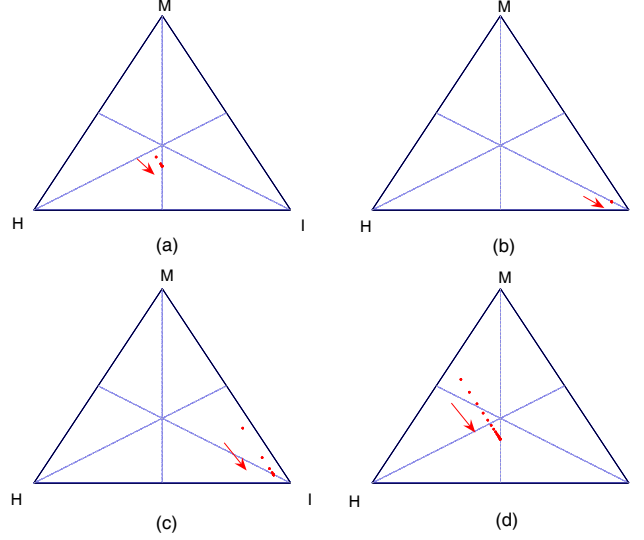


Figure 10: Effect of Information on Interactions.

response to the informational structure. The lower left vertex represents (**H**)istory, the total sum of weights assigned to last period's choices. The lower right vertex represents (**I**)nformation, the sum of weights on current types observed. Finally, the upper vertex represents average information, (**M**)ean type, $\bar{\theta}$. In part (a), we have *mild preferences* for conformity once again. The dots are concentrated near the middle of the triangle (equal weights on history, information, and mean type) and they do not move much as a response to changes in the amount of current information. Part (d) is the counterpart with *strong interactions*. Hence the significant change from almost no weight on current information to almost equal weights. Individuals use the information in the best possible way by putting more weight on it in their policy functions. This is due to the fact that forming expectations more precisely how the neighbors will behave becomes more important for each agent, due to the increased strength of interactions. Part (c) is *mild interactions* but *strong own-type effect* ($\frac{\alpha_1}{\alpha_2} = \frac{1}{20}$, $\frac{\alpha_2}{\alpha_3} = 20$) and part (b) is *strong interactions* and *strong own-type effect* ($\frac{\alpha_1}{\alpha_2} = \frac{1}{20}$, $\frac{\alpha_2}{\alpha_3} = 1$). We do not see much change in (b), although most of the total weight is put on information. This is mainly due to the fact that any agent a cares so much about his current type that, he neglects the other effects. In (c), although the own-effect is still strong, due to the strength of interactions, each agent uses the average information to form the best expectations regarding the behavior of the other agents. As the amount of information increases, each agent forms better expectations by transferring the policy weight from average information to precise information on close neighbors.

6 Identification

We study here the identification of social interactions in the context of our linear dynamic economy with conformity. Identification obtains when the restrictions imposed on actions at equilibrium by preferences for conformity are distinct from those imposed by other relevant structural models.⁴² Consider in particular an alternative structural model characterized by (cross-sectionally) correlated preferences across agents. This specific alternative model is focal because correlated preferences could be generally due to some sort of assortative matching or positive selection in social interaction, which induces agents with correlated preferences to interact socially. Suppose an econometrician observes panel data of individual actions over time displaying spatial correlation of individual actions at each time. Such correlation can generally be due to social interactions, as our analysis has shown. Such correlation could also ensue, however, from the spatial correlation of preference types, which we have excluded by assumption in our analysis to this point. But is there any structure in the spatial correlation which is implied by preference for conformity and not by correlated preferences? An affirmative answer to this question implies that the social interaction model is identified with respect to the correlated preferences model.

The structural analysis of identification in linear economies with social interactions starts with Manski (1993).⁴³ Manski restricts his analysis to static linear models, or, more specifically, linear economies in which the social interactions operate through the mean action in a pre-specified group, (*linear in means models*). In this context, identification is problematic due to the colinearity problem introduced by the mean action, the so-called *reflection problem*, and due to the possible correlation of unobservables. In the context of linear in means models, a recent literature has studied identification under the condition that the population of agents could be partitioned into a sequence of finitely-populated non-overlapping groups; see e.g., Graham and Hahn (2005).⁴⁴

The economies we study in this paper are related to those studied by Manski (1993) and others in that we maintain linearity, an assumption which renders identification harder (see Blume et

⁴²The question of identification in economics has been clearly defined by Koopmans (1949) and Koopmans and Reiersøl (1950). The issue of identification goes back to Pigou (1910), Schultz (1938), Frisch (1928, 1933, 1934, and 1938), and Frisch et al (1931). By *identification* we mean *identification in population* (Sometimes *identification in population* is called *identifiability*; see e.g., Chiappori and Ekeland, 2009). See also Marschak (1942), Haavelmo (1944), Koopmans, Rubin, and Leipnik (1950), Wald (1950), Hurwicz (1950). More recent surveys on the topic exist of course; see Rothenberg (1971), Hausman and Taylor (1983), Hsiao (1983), Matzkin (2007), and Dufour and Hsiao (2008).

⁴³Blume et al. (2011), Blume and Durlauf (2005), Brock and Durlauf (2007), Graham (2011), and Manski (1993, 2000, 2007) survey the main questions pertaining to identification in this social context.

⁴⁴Also: Davezies, D'Haultfoeuille and Fougère (2009) extends these results exploiting variation over the size of the populations; Graham (2008) uses excess variance across groups; Bramoullé et al (2009) uses reference group heterogeneity for identification. Other recent contributions include Glaeser and Scheinkman (2001), Graham and Hahn (2005); De Paula (2009), Evans, Oates and Schwab (1992), Ioannides and Zabel (2008), and Zanella (2007).

al. (2011)). On the other hand we introduce several fundamental distinguishing features: in particular, we allow for more general forms of social interactions across agents and for dynamic economies. More precisely, in the class of economies we study, the equilibrium action of agent a in an infinite horizon economy satisfies

$$x_t^a = \beta_1 x_{t-1}^a + \beta_2 \theta_t^a + \sum_{b \neq 0} \beta_{3,b} x_t^{a+b};$$

while in a linear in means economy the corresponding equation is:⁴⁵

$$x^a = \beta \theta^a + \gamma \sum_{b \in N(a)} x^b.$$

By studying populations composed of an infinite number of overlapping neighborhoods our analysis sheds some light on the nature of identification results which exploit an infinite number of non-overlapping groups, as in Graham and Hahn (2005) and in the literature discussed in footnote 44. The overlapping structure of our neighborhoods, in fact, breaks the independence which is required when non-overlapping groups are considered.⁴⁶ Furthermore, by studying dynamic models we are able to exploit the theoretical implications deriving from the optimality of the dynamic choices of agents on time series autocorrelations of actions, over and above the implications regarding the cross-sectional (spatial) correlations. In a related context, de Paula (2009) and Brock and Durlauf (2010) also exploit the properties of dynamic equilibrium, the discontinuity in adoption curves in their continuous time model, to identify social interactions.⁴⁷

We turn to our main identification results. The first series of results regards the identification of the dynamic structure - that is, distinguishing the properties of dynamic social interaction economies from those of myopic (hence static) economies. The second series of results regards instead the identification of social interactions, that is, distinguishing preference for conformity from correlated preferences.

6.1 Dynamic Rationality vs. Myopia

In this section we compare equilibrium configurations of dynamic economies with rational agents with those of economies with myopic agents. When agents are myopic, even economies with a dynamic structure, e.g., when agents' actions at time t depend on their previous actions, are

⁴⁵Note that, to ease the comparison we adopt here the best-reply representation of equilibrium actions; see equation (5).

⁴⁶We maintain however the assumption of symmetric neighborhoods, an assumption which, as is the case for linearity, renders identification harder: see Bramoullé, Djebbari and Fortin (2009) for a study of the identification power of observable asymmetric neighborhoods.

⁴⁷See also Cabral (1990) for an early discussion of these issues and Young (2009); see Blume, Brock, Durlauf, Ioannides (2011) for an up to date survey.

effectively static. These economies have been extensively studied in the theoretical and empirical literature on social interactions, following the mathematical physics literature in statistical mechanics on interacting particle systems. Suppose that myopic agents, when called to make a choice, act as if they expect never to be called to act again.⁴⁸ Given initial history x_{t-1} and realization θ_t , each myopic agent $a \in \mathbb{A}$ chooses $x_t^a \in X$ to maximize

$$u(x_{t-1}^a, x_t^a, x_t^{a-1}, x_t^{a+1}, \theta_t^a) := -\alpha_1(x_{t-1}^a - x_t^a)^2 - \alpha_2(\theta_t^a - x_t^a)^2 - \alpha_3(x_t^{a-1} - x_t^a)^2 - \alpha_3(x_t^{a+1} - x_t^a)^2$$

There exists then a unique symmetric policy function for any agent a , $g^{a,m}$ (m for myopic) such that

$$g^{a,m}(x_{t-1}, \theta_t; \alpha) := \sum_{b \in \mathbb{A}} c^{b,m} x_{t-1}^{a+b} + \sum_{b \in \mathbb{A}} d^{b,m} \theta_t^{a+b}$$

where we make explicit the dependence of the policy function on the preference parameters $\alpha = (\alpha_1, \alpha_2, \alpha_3)$.⁴⁹ The coefficients of the policy function $g^{a,m}$ are equal to the ones of the unique MPE policy function of a one-period ($T = 1$) social interactions economy: $c^{b,m} = c_1^b$, $d^{b,m} = d_1^b$, for $b \in \mathbb{A}$. In this sense, myopic models are nested within the class of dynamic models we study.

In the following we ask whether the spatial correlations generated by the long-run stationary distribution of an infinite-horizon model can be distinguished from those obtained as the limit distribution of a myopic model. Let $g^a(x_{t-1}, \theta_t^a; \alpha)$ denote agent a 's policy function from the dynamic social interaction model, where we make once again explicit the dependence of the policy function on α . We say that $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$ is a stochastic process induced by the dynamic economy with parameters α if it satisfies

$$x_t^a = g^a(x_{t-1}, \theta_t^a; \alpha), \text{ for any } a \in \mathbb{A} \text{ and any } t \geq 1$$

We instead say that $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$ is a stochastic process induced by the myopic economy with parameters α if it satisfies

$$x_t^a = g^{a,m}(x_{t-1}, \theta_t; \alpha), \text{ for any } a \in \mathbb{A} \text{ and any } t \geq 1$$

⁴⁸See e.g., Blume and Durlauf (2001), Brock and Durlauf (2001b); and Glaeser and Scheinkman (2003) for a comprehensive survey. Liggett (1985) is the standard reference for the mathematical literature.

⁴⁹In some of the literature, myopic agents are modelled not only as assuming that all agents in the economy only interact once, but also that their neighbors are not changing their previous period actions. In this case an agent a solves

$$\max_{x_t^a \in X} -\alpha_1(x_{t-1}^a - x_t^a)^2 - \alpha_2(\theta_t^a - x_t^a)^2 - \alpha_3(x_{t-1}^{a-1} - x_t^a)^2 - \alpha_3(x_{t-1}^{a+1} - x_t^a)^2.$$

and his policy function is

$$x_t^a = \beta_1 x_{t-1}^a + \beta_2 \theta_t^0 + \beta_3 x_{t-1}^{-1} + \beta_3 x_{t-1}^1.$$

It can be shown, see Glaeser and Scheinkman (2003), that the ergodic stationary distribution of actions in this economy coincides with that of *myopic* agents as defined in the text. As a consequence, our identification results extend to this economy as well.

We are now ready to introduce our definition of identification of social interactions.

Definition 3 Let $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$ denote a stochastic process induced by the dynamic economy with parameters α . We say that the dynamic economy with parameters α is identified with respect to myopic economies if there does not exist an $\hat{\alpha}$, such that the process $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$ is also induced by a myopic economy with parameter $\hat{\alpha}$.

The characterization of the spatial correlation of actions at equilibrium for different time-horizons T , which we provided in Section 4.2, gives us a straightforward answer to the identification question. Recall in fact that the coefficients of the policy function $g^{a,m}$ are equal to the ones of the unique MPE policy function of a one-period ($T = 1$) social interactions economy. Recall also that the covariances between agents's choices obtained from data generated by a typical model of infinite-horizon stationary social interactions are fundamentally different from those generated by a myopic model. In particular, we have shown in Section 4.2 that, for a typical choice of α ,

$$r_{a,T} = \frac{\rho_{a+1,T}}{\rho_{a,T}}$$

is non-monotonic in a , for longer horizon economies; and so is r_a , the ratio of the limit economy with $T = \infty$; while $r_{a,1}$ declines monotonically in a , for any α ; see Figure 4. Moreover, the limit unconditional covariances inherit the (non)-monotonicity features of their one-step conditional counterparts. Finally, by continuity, the non-monotonicity property necessarily holds for an open set of the parameter space, and is hence *robust*. Summarizing, then, we have the following.

Proposition 2 (Rationality vs. Myopia) *A dynamic economy with parameter α is identified with respect to myopic economies, for a robust subset α .*

As an illustration we present in Figure 11, r_a as a function of a , at the stationary distribution, for different levels of strength of interaction proxied by the ratio $\left(\frac{2\alpha_3}{\Delta_1}\right)$.⁵⁰ Clearly, for a large set of parameters, non-monotonicity obtains at the stationary distribution. The limit auto-correlation function for the myopic model, on the contrary, inherits the behavior of its one-step transition counterpart: it converges at a monotonically decreasing rate.

Consider an econometrician fitting a static (myopic) model through data generated by the dynamic equilibrium of an economy with parameter α . From Proposition 1, $r_1(\alpha) < r(\alpha)$ for any possible α . As a consequence, the parameter $\hat{\alpha}$ estimated by the econometrician imposing the static (myopic) structure on the data, will satisfy

$$r_1(\hat{\alpha}) = r(\alpha) > r_1(\alpha) \tag{17}$$

⁵⁰More precisely, we set $\left(\frac{\alpha_1}{\alpha_2}\right) = 1$, $\left(\frac{2\alpha_3}{\Delta_1}\right) \in \{0.1, 0.2, 0.5, 0.75, 0.9\}$ and $\beta = .95$.

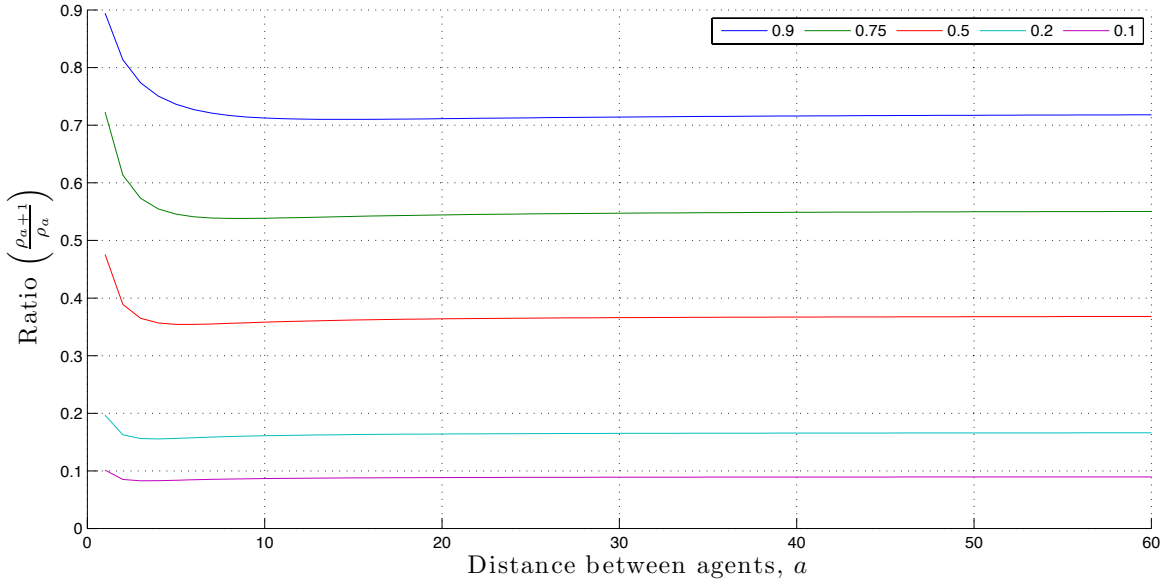


Figure 11: The ratio r_a as a function of a for different values of $\left(\frac{2\alpha_3}{\Delta_1}\right)$ at the stationary distribution.

From (51), however, r_1 is monotonically decreasing in $\left(\frac{\Delta_1}{2\alpha_3}\right)$. As a consequence,

$$\left(\frac{\hat{\alpha}_3}{\hat{\alpha}_1 + \hat{\alpha}_2 + 2\hat{\alpha}_3}\right) > \left(\frac{\alpha_3}{\alpha_1 + \alpha_2 + 2\alpha_3}\right),$$

and the econometrician **overestimates** the social interaction effects.

6.2 Social Interactions vs. Selection

In our dynamic economies, spatial correlation of individual actions at each time is induced by social interactions and preference for conformity. But spatial correlation of actions could be induced in principle also by spatial correlation of preference types, with no social interaction. Take two agents, e.g., agent a and agent b . A positive correlation between x_t^a and x_t^b could be due to a positive correlation between θ_t^a and θ_t^b . In this last case, preferences for conformity and social interactions would play no role in the correlation of actions at equilibrium. Rather, such correlation would be due to the fact that agents have correlated preferences. As we already noted, correlated preferences could be generally due to some sort of assortative matching or positive selection in social interaction, which induces agents with correlated preferences to interact socially.

In our economy, at a symmetric Markov perfect equilibrium, each agent $a \in \mathbb{A}$ acts according to the policy function $g_{T-(t-1)}^a(x_{t-1}, \theta_t^a; \alpha)$, where we make once again explicit the dependence of

the policy function on the preference parameters $\alpha = (\alpha_1, \alpha_2, \alpha_3)$. If $T = \infty$, the policy function is stationary $g^a(x_{t-1}, \theta_t^a; \alpha)$. Recall that the parameter α_3 represents the weight of conformity in each agent's preferences. It follows that $\alpha_3 = 0$ corresponds to an economy with no social interactions. We say that $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$ is a stochastic process induced by α and $(\theta_t^a)_{t \geq 1}^{a \in \mathbb{A}}$ if it satisfies

$$x_t^a = g_{T-(t-1)}^a(x_{t-1}, \theta_t^a; \alpha), \text{ for any } a \in \mathbb{A} \text{ and any } T \geq t \geq 1$$

We are now ready to construct our definition of identification of social interactions.

Definition 4 Let $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$ denote a stochastic process induced by α and $(\theta_t^a)_{t \geq 1}^{a \in \mathbb{A}}$, where $(\theta_t^a)_{t \geq 1}^{a \in \mathbb{A}}$ is i.i.d. across agents and serially uncorrelated, that is, where $\text{Cov}(\theta_t^a, \theta_t^b) = \text{Cov}(\theta_t^a, \theta_{t+1}^a) = 0$ for any $a \neq b \in \mathbb{A}$ and any $t \geq 0$. We say that α is identified if there does not exist an $\hat{\alpha}$, with $\hat{\alpha}_3 = 0$, such that the process $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$ is also induced by $\hat{\alpha}$ and some stochastic process $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$. We say that social interactions are identified if some α , with $\alpha_3 > 0$, is identified.

The conditions for identification of social interactions can be weakened by restricting the stochastic process $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$. We say that α is (resp. social interactions are) *identified relatively to a set of preference shocks* if $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ in Definition 4 is required to belong to a set of preference shocks which satisfies some specific restriction.

Finally, the conditions for identification of social interactions can be strengthened by limiting the observable properties of the process $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$.

We first consider the case of an infinite horizon economy: policy functions are stationary and an ergodic distribution exists. In this context, we study first the possibility of obtaining identification by observing the properties of the stationary distribution of actions rather than the whole panel $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$. We then pass on to identification tout court, that is exploiting the whole dynamic restrictions imposed by the model on $(x_t^a)_{t \geq 1}^{a \in \mathbb{A}}$, not just the restrictions on the stationary distribution. We shall see that results are negative in both cases, that is, identification is not obtained in general. Secondly, we study identification relatively to a series of relevant restrictions on the stochastic process for preference shocks $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$. These restrictions are meant to capture natural properties of the selection mechanism which induces agents to display spatially correlated preferences.

6.2.1 Infinite horizon (stationary) economies

Consider first the stationary distribution of actions as identified by its implied spatial correlation function ρ_b .

Proposition 3 *Social interactions are not identified by the properties of the spatial correlation function ρ_b of the stationary distribution of actions in infinite horizon economies.*

The proof is simple and instructive and hence it is reported following in the text.

Proof: We have shown in Section 3.3 that the stationary distribution of our dynamic economy with social interactions, that is, $\alpha_3 > 0$, and i.i.d. preference shock process $\{\theta_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$, is given by the ergodic measure π in (10), i.e. π is the joint distribution of

$$x_t = \left(\frac{e(\alpha) \bar{\theta}}{1 - C(\alpha)} + \sum_{s=1}^{\infty} \sum_{b_1} \cdots \sum_{b_s} c(\alpha)^{b_1} \cdots c(\alpha)^{b_{s-1}} \left(d(\alpha)^{b_s} \theta_{t+1-s}^{a+b_1+\cdots+b_s} \right) \right)_{a \in \mathbb{A}}$$

Consider now an alternative specification of our economy with *no interactions* between agents ($\hat{\alpha}_3 = 0$) and *no habits* ($\hat{\alpha}_1 = 0$) but simply a preference shock process $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ and *own type effects* with $\hat{\alpha}_2 > 0$. For this economy, equilibrium choice of agent a at time t is given by

$$x_t^a = \hat{\theta}_t^a$$

As long as the process $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ is the one where

$$\hat{\theta}_t^a := \frac{e(\alpha) \bar{\theta}}{1 - C(\alpha)} + \sum_{s=1}^{\infty} \sum_{b_1} \cdots \sum_{b_s} c(\alpha)^{b_1} \cdots c(\alpha)^{b_{s-1}} \left(d(\alpha)^{b_s} \theta_{t+1-s}^{a+b_1+\cdots+b_s} \right)$$

the probability distributions that the two specifications (with and without interactions) generate on the observables of interest, $\{x_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$, are identical. Hence, one cannot identify from the stationary distribution of choices which specification generates the data. $\blacksquare$

More generally, we investigate if the dynamic equilibrium restrictions of our model are sufficient to identify social interactions.

Proposition 4 *Social interactions are not identified in infinite horizon economies.*

This proof is also simple and instructive and hence it is reported following in the text.

Proof: In the case of complete information, the policy function is:

$$x_t^a = \sum_{b \in \mathbb{A}} c^b(\alpha) x_{t-1}^{a+b} + \sum_{b \in \mathbb{A}} d^b(\alpha) \theta_t^{a+b} + e \bar{\theta}$$

As we saw in Lemma 2, one can obtain by iteration the reduced form⁵¹

$$\begin{aligned} x_t^a &= \sum_{b_1} \cdots \sum_{b_t} c(\alpha)^{b_1} \cdots c(\alpha)^{b_t} x_0^{a+b_1+\cdots+b_t} \\ &\quad + \sum_{s=1}^t \sum_{b_1} \cdots \sum_{b_{s-1}} c(\alpha)^{b_1} \cdots c(\alpha)^{b_{s-1}} \left(\sum_{b_s} d(\alpha)^{b_s} \theta_{t-(s-1)}^{a+b_1+\cdots+b_s} + e(\alpha) \bar{\theta} \right) \end{aligned}$$

⁵¹In Lemma 2, the iteration stops once it reaches period 1. But, since a stationary MPE exists by Theorem 1, we iterate here once more on the form in Lemma 2 using the stationary policy function and write equilibrium choices as a function of initial conditions x_0 .

Consider now the alternative specification with *no interactions* between agents ($\hat{\alpha}_3 = 0$) and *no habits* ($\hat{\alpha}_1 = 0$), a preference shock process $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ and *own type effects* with $\hat{\alpha}_2 > 0$. For this economy, equilibrium choice of agent a at time t is given by

$$x_t^a = \hat{\theta}_t^a$$

Defining the new preference shock process $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ as

$$\begin{aligned} \hat{\theta}_t^a : &= \sum_{b_1} \cdots \sum_{b_t} c(\alpha)^{b_1} \cdots c(\alpha)^{b_t} x_0^{a+b_1+\cdots+b_t} \\ &+ \sum_{s=1}^t \sum_{b_1} \cdots \sum_{b_s} c(\alpha)^{b_1} \cdots c(\alpha)^{b_{s-1}} \left(d(\alpha)^{b_s} \theta_{t-(s-1)}^{a+b_1+\cdots+b_s} + e(\alpha) \bar{\theta} \right) \end{aligned}$$

would imply that for an arbitrary initial distribution π_0 for x_0 , the joint probability distributions that the two specifications (with and without interactions)) generate on the process $\{x_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$, are identical. Moreover, if one allows for infinite histories, one can define the preference shock process $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ as before

$$\hat{\theta}_t^a := \frac{e(\alpha) \bar{\theta}}{1 - C(\alpha)} + \sum_{s=1}^{\infty} \sum_{b_1} \cdots \sum_{b_s} (c(\alpha))^{b_1} \cdots (c(\alpha))^{b_{s-1}} \left((d(\alpha))^{b_s} \theta_{t-(s-1)}^{a+b_1+\cdots+b_s} \right)$$

and obtain observational equivalence once again. Hence, we conclude that identification is not possible. ■

An intuition about this result can be obtained by loosely reducing the identification of social interactions in infinite horizon economies to the well known problem of distinguishing a VAR from an MA(∞) process. Stacking in a vector $\mathbf{x}_t$ (resp. θ_t) the actions x_t^a over the index $a \in \mathbb{A}$ (resp. the preference shocks θ_t^a), policy functions can be loosely written as a VAR:

$$\mathbf{x}_t = \Phi \mathbf{x}_{t-1} + \delta_t, \quad \text{with } \delta_t = \Gamma \theta_t + e \bar{\theta}$$

where $E(\delta_t \delta_{t-\tau}) = 0$ for all $\tau > 0$. Under standard stationarity assumptions, the VAR has an MA(∞) representation

$$\mathbf{x}_t = (I_A - \Phi \mathbf{L})^{-1} \delta_t = \delta_t + \Psi_1 \delta_{t-1} + \Psi_2 \delta_{t-2} + \dots$$

for a sequence $\Psi_1, \Psi_2, \dots$ such that $(I_A - \Phi \mathbf{L})(I_A + \Psi_1 \mathbf{L} + \Psi_2 \mathbf{L}^2 + \dots) = I_A$. The argument in the proof of Proposition 4 therefore amounts to picking

$$\mathbf{x}_t = \hat{\theta}_t = \delta_t + \Psi_1 \delta_{t-1} + \Psi_2 \delta_{t-2} + \dots$$

6.2.2 Finite-horizon (non-stationary) economies

Consider now the case of a finite horizon economy. In this case the unique policy function and the distribution of actions are *non-stationary*, as we have shown, and hence identification might obtain in those environments where correlated effects satisfy a stationary law through time. Consider then a restriction to the class of admissible preference shock processes $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ which satisfy the following *conditional covariance stationarity* restriction:

Definition 5 (Conditional Covariance Stationarity) *A process $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ is said to be conditional covariance stationary if $\text{Cov}(\hat{\theta}_t^a, \hat{\theta}_t^b \mid \hat{\theta}_1, \dots, \hat{\theta}_{t-1}) = Z(a, b, \hat{\theta}_{t-n}, \dots, \hat{\theta}_{t-1}) \in \mathbb{R}$, for $a, b \in \mathbb{A}$, $t = n + 1, \dots, T$.⁵²*

This condition defines a large class of stochastic processes for which $\text{Cov}(\hat{\theta}_t^a, \hat{\theta}_t^b \mid \hat{\theta}_1, \dots, \hat{\theta}_{t-1}) = Z(a, b, \hat{\theta}_{t-n}, \dots, \hat{\theta}_{t-1}) \in \mathbb{R}$ depends on the position of agents a and b and on their finite memory (at most n) of realizations of the process. It is a relatively weak and natural condition in that it allows for the intertemporal dependence to be a function of the position of the agents; what it excludes is events in the distant past from having a significant effect on the *joint* determination of agents' types today.⁵³

Conditional covariance stationarity of preference shocks is in fact sufficient for identification of social interactions.

Proposition 5 *Social interactions are identified relatively to processes satisfying the conditional covariance stationarity restriction with $n \leq T - 2$.*

While the spatial autocorrelations between agents' choices are the same across periods in the absence of interaction effects, they vary in presence of social interactions.⁵⁴ This non-stationarity feature of the equilibrium process is at the heart of the proof, whose details follow.

⁵²The original definition is due to Mandelbrot (1967) who provides the conditional spectral analysis of *sporadically varying* random functions in the mathematical theory of information transmission with noise. In his environment, he requires $E[\hat{\theta}_t \hat{\theta}_{t+n} \mid 1 \leq t < t+n \leq T]$ to be independent of t . Ours is a slightly weaker condition since it uses fixed finite memory. For more recent usage of conditional covariance restrictions see the Times Series literature studying persistence of conditional variances (ARCH, GARCH), especially Bollerslev (1986), Bollerslev and Engle (1993), Engle (1982), Engle and Bollerslev (1986). For a survey of these models see Bauwens et al. (2006) and Bollerslev et al. (1992, 1994). For analogous conditions of weak (unconditional) stationarity used in the times series literature see Hamilton (1994), p. 45-46 and chapter 10.

⁵³All existing social interaction models we can think of have stochastic structures that are special cases of this class. More specifically, they typically assume either time-independent or finite memory Markov structures to model exogenous effects; see e.g., Brock and Durlauf (2001b), Conley and Topa (2003, 2007), de Paula (2009), Glaeser and Scheinkman (2001), Nakajima (2007), Topa (2001), and Young (2009).

⁵⁴More specifically, they can be ordered with respect to their spatial rate of tail convergence; see Proposition 1.

Proof: Consider a finite-horizon, T -period economy with $T \geq 2$. We assume that $n \leq T - 2$. In the absence of interactions ($\hat{\alpha}_3 = 0$), agent a 's final period optimal choice is⁵⁵

$$x_T^a = c_1(\hat{\alpha}) x_{T-1}^a + d_1(\hat{\alpha}) \hat{\theta}_T^a \quad (18)$$

Thanks to the linearity of the policy functions across periods with ($\hat{\alpha}_3 = 0$), any path of shock realizations $(\hat{\theta}_1, \dots, \hat{\theta}_{T-1})$, given x_0 , generates a path of configurations $(x_0, x_1, \dots, x_{T-1})$. Thus, conditioning on all imaginable choice paths spans all imaginable preference shock paths, given that the observables are generated by the above-mentioned policy functions. This observation is useful because the a -step covariance between equilibrium choices of agent 0 and a in case of *interactions* is given by

$$\begin{aligned} \text{Cov} \left(x_T^0, x_T^a \middle| x_{T-1} \right) &= \text{Cov} \left(x_T^0, x_T^a \middle| x_0, x_1, \dots, x_{T-1} \right), \quad \forall (x_0, x_1, \dots, x_{T-1}) \\ &= \text{Cov} \left(\sum_{b_1 \in \mathbb{A}} d_1^{b_1} \theta_T^{b_1}, \sum_{b_2 \in \mathbb{A}} d_1^{b_2} \theta_T^{a+b_2} \right) \\ &= \text{Var}(\theta) \sum_{b_1 \in \mathbb{A}} d_1^{b_1} d_1^{b_1-a} \end{aligned} \quad (19)$$

which means that the covariance term is independent of the conditioned upon path. The implication of this is that in order the specification with no interactions to be observationally indistinguishable from the interactions case, the a -step conditional covariances, computed using (18)

$$\begin{aligned} \text{Cov} \left(x_T^0, x_T^a \middle| x_{T-1} \right) &= d_1(\hat{\alpha})^2 \text{Cov} \left(\hat{\theta}_T^0, \hat{\theta}_T^a \middle| x_0, \hat{\theta}_1, \dots, \hat{\theta}_{T-1} \right) \\ &= d_1(\hat{\alpha})^2 Z(0, a, \hat{\theta}_{T-n}, \dots, \hat{\theta}_{T-1}), \quad \forall (x_0, \hat{\theta}_1, \dots, \hat{\theta}_{T-1}) \\ &= d_1(\hat{\alpha})^2 \bar{Z}(0, a) \in \mathbb{R}, \quad \forall (\hat{\theta}_{T-n}, \dots, \hat{\theta}_{T-1}) \end{aligned} \quad (20)$$

should be independent of previous realizations and vary only with respect to the positions of the agents in the network and not with respect to previous realizations. The function $\bar{Z}$ is implicitly defined to capture that fact and so that the expression in (20) matches the one in (19). This is not an assumption but an equilibrium restriction and we are bound to choose the covariance structure accordingly.

Once again, in order the two model to be observationally indistinguishable, the a -step conditional covariances, for all $a \in \mathbb{A}$, under interactions and in the absence of interactions should

⁵⁵ In the absence of interactions, problems (23) and (28) become individual maximization problems. Elementary dynamic programming techniques, as in Stokey and Lucas (1989), yield the policy functions we use here.

match, in period $T - 1$. Similar calculations in period $T - 1$ yield

$$\begin{aligned}
Cov\left(x_{T-1}^0, x_{T-1}^a \middle| x_{T-2}\right) &= Cov\left(x_{T-1}^0, x_{T-1}^a \middle| x_0, x_1, \dots, x_{T-2}\right), \quad \forall (x_0, x_1, \dots, x_{T-2}) \\
&= Var(\theta) \sum_{b_1 \in \mathbb{A}} d_2^{b_1} d_2^{b_1-a} \\
&= d_2(\hat{\alpha})^2 Cov\left(\hat{\theta}_{T-1}^0, \hat{\theta}_{T-1}^a \middle| x_0, \hat{\theta}_1, \dots, \hat{\theta}_{T-2}\right), \quad \forall (x_0, \hat{\theta}_1, \dots, \hat{\theta}_{T-2}) \\
&= d_2(\hat{\alpha})^2 Z(0, a, \hat{\theta}_{T-1-n}, \dots, \hat{\theta}_{T-2}) \in \mathbb{R}, \quad \forall (\hat{\theta}_{T-1-n}, \dots, \hat{\theta}_{T-2}) \\
&= d_2(\hat{\alpha})^2 \bar{Z}(0, a), \quad \forall (\hat{\theta}_{T-1-n}, \dots, \hat{\theta}_{T-2})
\end{aligned} \tag{21}$$

where the first two equalities are as in (19); third is the restriction imposed by observable indistinguishability; fourth by conditional covariance stationarity; finally fifth is by conditional covariance stationarity across periods using (20). Putting the equilibrium restrictions in periods $T - 1$ and T together, using (20) and (21), we can write

$$\begin{aligned}
\frac{Cov\left(x_{T-1}^0, x_{T-1}^{a+1} \middle| x_{T-2}\right)}{Cov\left(x_{T-1}^0, x_{T-1}^a \middle| x_{T-2}\right)} &= \left(\frac{d_2(\hat{\alpha})^2}{d_2(\hat{\alpha})^2}\right) \left(\frac{\bar{Z}(0, a+1)}{\bar{Z}(0, a)}\right) \\
&= \left(\frac{d_1(\hat{\alpha})^2}{d_1(\hat{\alpha})^2}\right) \left(\frac{\bar{Z}(0, a+1)}{\bar{Z}(0, a)}\right) \\
&= \frac{Cov\left(x_T^0, x_T^{a+1} \middle| x_{T-1}\right)}{Cov\left(x_T^0, x_T^a \middle| x_{T-1}\right)}
\end{aligned}$$

Since the choice of a is arbitrary, we can look at the same expression as a becomes progressively larger. So, as $a \rightarrow \infty$, the expression should give

$$\lim_{a \rightarrow \infty} \frac{Cov\left(x_{T-1}^0, x_{T-1}^{a+1} \middle| x_{T-2}\right)}{Cov\left(x_{T-1}^0, x_{T-1}^a \middle| x_{T-2}\right)} = r_2 = r_1 = \lim_{a \rightarrow \infty} \frac{Cov\left(x_T^0, x_T^{a+1} \middle| x_{T-1}\right)}{Cov\left(x_T^0, x_T^a \middle| x_{T-1}\right)} \tag{22}$$

which is a contradiction to Proposition 1. Therefore, there does not exist a conditional covariance stationary preference shock process $\{\hat{\theta}_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ that generates an equilibrium choice process $\{x_t^a\}_{t \geq 1}^{a \in \mathbb{A}}$ under the no interactions specification ($\hat{\alpha}_3 = 0$) that is observationally equivalent to the process generated by the local interactions ($\alpha_3 \neq 0$) process. This concludes the proof. $\blacksquare$

7 Extensions

The class of social interaction economies we studied in this paper has been restricted along several dimensions to better provide a stark theoretical analysis. Some of these restrictions, however, turn out to be important in applications and empirical work. In this section, therefore, we illustrate

how our analysis can be extended to study more general neighborhood network structures for social interactions, more general stochastic processes for preference shocks, the addition of global interactions, that is, interactions at the population level, and the effects of stock variables which carry habit effects.

7.1 General Neighborhood Network Structures

Throughout the paper, we studied symmetric neighborhood structures. This is generalized easily. Consider an arbitrary neighborhood network structure (not necessarily translation invariant), $N : \mathbb{A} \rightarrow 2^{\mathbb{A}}$. Suppose also that a generic agent a 's preferences are represented by the utility function u^a defined as

$$u^a \left(x_{t-1}^a, x_t^a, \{x_t^b\}_{b \in N(a)}, \theta_t^a \right) := -\alpha_{a,1} (x_{t-1}^a - x_t^a)^2 - \alpha_{a,2} (\theta_t^a - x_t^a)^2 - \sum_{b \in N(a)} \alpha_{a,b} (x_t^b - x_t^a)^2$$

Notice that we allow for the preferences of any two agents a and b to be arbitrarily different in their parametrization, provided either $\alpha_{a,1} > 0$ or $\alpha_{a,2} > 0$ and $\sum_{b \in N(a)} \alpha_{a,b} < \infty$ so that peer effects are bounded. Under this specification, best-responses are well defined, interior, and well-behaved. An MPE exists and the policy function of an arbitrary agent $a \in \mathbb{A}$ at equilibrium has the following form

$$g_{T-(t-1)}^a(x_{t-1}, \theta_t) = \sum_{b \in \mathbb{A}} c_{T-(t-1)}^{a,b} x_t^b + \sum_{b \in \mathbb{A}} d_{T-(t-1)}^{a,b} \theta_t^b + e_{T-(t-1)}^a \bar{\theta}$$

where, as before, all coefficients are non-negative and $\sum_{b \in \mathbb{A}} (c_t^{a,b} + d_t^{a,b}) + e_t^a = 1$. For uniqueness of equilibrium, it is sufficient that the relative composition of the peer effects within the determinants of individual choice be uniformly bounded, i.e., that there exists a positive constant K such that for each individual $a \in \mathbb{A}$

$$\frac{\sum_{b \in N(a)} \alpha_{a,b}}{\alpha_{a,1} + \alpha_{a,2} + \sum_{b \in N(a)} \alpha_{a,b}} < K.$$

Under this condition,⁵⁶ best responses induce a contraction operator and we obtain a unique equilibrium for any finite-horizon economy.

Ergodicity (relative to a given MPE) and welfare results extend straightforwardly, as do identification results. Notably, our positive identification result for non-stationary economies, Proposition 5, also extends: since preference parameters of any agent a are stationary, in a finite-horizon economy, correlations of equilibrium actions between agents vary only due to interactions for preference processes that satisfy a Conditional Covariance Stationarity restriction.

⁵⁶A related condition is referred to, in the literature, as the *Moderate Social Influence* condition; see e.g. Glaeser and Scheinkman (2003) and Horst and Scheinkman (2006).

7.2 General Stochastic Processes for Preference Shocks

The agents in our model make their decisions based on past behavior and current shocks. Our analysis however extends straightforwardly to economies where shocks are persistent across time as long as the economy is one of complete information.⁵⁷ We give here, as an illustration, an example of Markov dependence, where at any given period the probability of next period shocks depends on current realizations.

Consider any T -period economy with $T \leq \infty$. Recall from Section 2 that preference shocks $\theta_t := (\theta_t^a)_{a \in \mathbb{A}}$ are defined on the canonical probability space $(\Theta, \mathcal{F}, \mathbb{P})$, where $\Theta := \{(\theta^a)_{a \in \mathbb{A}} : \theta^a \in \Theta\}$. Let $Q : \Theta \times \mathcal{F} \rightarrow \mathbb{R}_+$ be a *transition function* such that

- (i) for any period t and any $\theta \in \Theta$, $Q(\theta, A) = \Pr\{\theta_{t+1} \in A \mid \theta_t = \theta\}$, for all $A \in \mathcal{F}$.
- (ii) for each $A \in \mathcal{F}$, $Q(\cdot, A)$ is $\mathcal{F}$ -measurable.

Any agent $a \in \mathbb{A}$ solves the problem in (1) with persistent shocks where the expectation operator acts on the distribution induced by Q and other agents' strategies. We can write the problem recursively. The policy function of an arbitrary agent $a \in \mathbb{A}$ at equilibrium, in this economy, is

$$g_{T-(t-1)}^a(x_{t-1}, \theta_t) = \sum_{b \in \mathbb{A}} c_{T-(t-1)}^b x_{t-1}^{a+b} + \sum_{b \in \mathbb{A}} d_{T-(t-1)}^{a+b} \theta_t^a + e_{T-(t-1)}(\theta_t, a)$$

for some positive coefficients $(c_{T-(t-1)}^b)_{b \in \mathbb{A}}$, $(d_{T-(t-1)}^{a+b})_{b \in \mathbb{A}}$, and some constant $e_{T-(t-1)}(\theta_t, a)$ that depends only on the current type profile and on the agent's name, a .

Once again, existence, ergodicity (relative to a given MPE), and welfare results extend straightforwardly, as well as our identification results.

7.3 Global Interactions

Introducing global determinants of individual behavior into our framework is also relatively straightforward.⁵⁸ In particular, consider an economy in which the preferences of each agent $a \in \mathbb{A}$ depend also on the average action of the agents in the economy. Let the average action given a choice profile x be defined as

$$p(x) := \lim_{n \rightarrow \infty} \frac{1}{2n+1} \sum_{a=-n}^n x^a,$$

⁵⁷The mathematical issues arising in dynamic models with incomplete information are both well-known and outside the scope of the present paper. See Mailath and Samuelson (2006) for an extensive survey.

⁵⁸With respect to the analysis of MPE with local and global interactions in finite economies (as e.g., in Blume and Durlauf, 2001, and in Glaeser and Scheinkman, 2003), a few technical subtleties arise in our economy due to the infinite number of agents. The techniques we use are extensions of the ones we used in a previous paper, Bisin, Horst, and Özgür (2006). We refer the reader to this paper for details. Some of the needed mathematical analysis is developed in Föllmer and Horst (2001) and Horst and Scheinkman (2006).

when the limit exists. Let $\mathbf{X}_e$ denote the set of all configurations such that the associated average action exists:

$$\mathbf{X}_e := \left\{ x \in \mathbf{X} : \exists p(x) := \lim_{n \rightarrow \infty} \frac{1}{2n+1} \sum_{a=-n}^n x^a \right\}$$

The preferences of the agent $a \in \mathbb{A}$ in period t are described by the instantaneous utility function $u : \mathbf{X}_e \times \Theta \rightarrow \mathbb{R}$ of the conformity class

$$\begin{aligned} u^a \left(x_{t-1}^a, x_t^a, \{x_t^b\}_{b \in N(a)}, \theta_t^a, p(x_t) \right) := \\ -\alpha_{a,1} (x_{t-1}^a - x_t^a)^2 - \alpha_{a,2} (\theta_t^a - x_t^a)^2 - \alpha_{a,3} (p(x_t) - x_t^a)^2 - \sum_{b \in N(a)} \alpha_{a,b} (x_t^b - x_t^a)^2 \end{aligned}$$

Given $x \in \mathbf{X}_e$, the initial configuration of actions, a symmetric stationary MPE of a dynamic economy with local and global interactions is a map $g : \mathbf{X} \times \Theta \times X \rightarrow X$ and a map $F : X \rightarrow X$ such that:

$$\begin{aligned} g(x_{t-1}, \theta_t, p_t) &= \arg \max_{x_t^0 \in X} E \left[u \left(x_{t-1}^0, x_t^0, \left\{ g \left(R^b x_{t-1}, R^b \theta_t, p_t \right) \right\}_{b \in N(0)}, \theta_t^0, p_t \right) \right. \\ &\quad \left. + \beta V_g \left(x_t^0, \left\{ g \left(R^b x_{t-1}, R^b \theta_t, p_t \right) \right\}_{b \neq 0}, \theta_{t+1}, p_{t+1} \right) \mid (x_{t-1}, \theta_t) \right] \end{aligned}$$

and

$$p_{t+1} = F(p_t),$$

and

$$p_1 = p(x) \quad \text{and} \quad p_t = p(x_t) \quad \text{almost surely.}^{59}$$

At a symmetric MPE, any agent rationally anticipates that all others play according to the policy function g and also anticipates the sequence of average actions $\{p(x_t)\}_{t \in \mathbb{N}}$ to be determined recursively via the map F .

For this economy, we can show that the endogenous sequence of average actions $\{p(x_t)\}_{t \in \mathbb{N}}$ exists almost surely if the initial configuration x belongs to $\mathbf{X}_e$, and that it follows a deterministic recursive relation.⁶⁰ As a consequence, our main results extend and the policy function of an arbitrary agent $a \in \mathbb{A}$ at equilibrium is

$$g^a(x_{t-1}, \theta_t) = \sum_{b \in \mathbb{A}} c_b x_{t-1}^{a+b} + \sum_{b \in \mathbb{A}} d_b \theta_t^{a+b} + e \bar{\theta} + B^*(p(x_0))$$

for some positive coefficients $(c^b)_{b \in \mathbb{A}}$, $(d^b)_{b \in \mathbb{A}}$, e , and some constant $B^*(p(x_0))$ that depends only on the initial average action, $p(x_0)$.

⁵⁹As before, symmetry allows us to define a symmetric MPE with respect to agent 0. For an arbitrary agent $a \in \mathbb{A}$, then, his policy function is $g^a(x_{t-1}, \theta_t, p_t) = g(R^a x_{t-1}, R^a \theta_t, p_t)$, for all $(x_{t-1}, \theta_t, p_t) \in \mathbf{X} \times \Theta \times X$.

⁶⁰Linearity is crucial for these results. Only in this case, in fact, can the dynamics of average actions $\{p(x_t)\}_{t \in \mathbb{N}}$ be described recursively. In models with more general local interactions, the average action typically is not a sufficient statistic for the aggregate behavior of the configuration x ; hence a recursive relation typically fails to hold. In such more general cases, the analysis must be pursued in terms of empirical fields. Interested reader should consult Föllmer and Horst (2001).

7.4 Social Accumulation of Habits

In this section, we generalize the class of the economies we have studied to encompass a richer structure of dynamic dependence of agents' actions at equilibrium. Consider an economy where preferences of agent $a \in \mathbb{A}$ are represented by a utility function

$$u\left(S_t^a, x_t^a, \{x_t^b\}_{b \in N(a)}, \theta_t^a\right) := -\alpha_1 (S_t^a - x_t^a)^2 - \alpha_2 (\theta_t^a - x_t^a)^2 - \sum_{b \in N(a)} \alpha_{a,b} (x_t^b - x_t^a)^2$$

where S_t^a represents an accumulated stock variable,

$$S_{t+1}^a = (1 - \delta) S_t^a + x_t^a$$

For instance, S_t^a captures what the addiction literature calls a “*reinforcement effect*” on agent a 's substance consumption. In this economy the policy function at equilibrium is

$$g_{T-(t-1)}^a(S_t, \theta_t) = \sum_{b \in \mathbb{A}} c_{T-(t-1)}^{a,b} S_t^b + \sum_{b \in \mathbb{A}} d_{T-(t-1)}^{a,b} \theta_t^b + e_{T-(t-1)}^a \bar{\theta}$$

Note that in equilibrium each agent's choice depends on the stock of his neighbors' actions, that is, on their long-term behavioral patterns rather than just their previous period actions. Also, as the final period approaches, agent a assigns *uniformly* higher weights to his own stock.

8 Conclusion

Social interactions provide a rationale for several important phenomena at the intersection of economics and sociology. The theoretical and empirical study of economies with social interactions, however, has been hindered by several obstacles. Theoretically, the analysis of equilibria in these economies induces generally intractable mathematical problems: equilibria are represented formally by a fixed point in configuration of actions, typically an infinite dimensional object; and embedding equilibria in a full dynamic economy adds a second infinite dimensional element to the analysis. Computationally, these economies are also generally plagued by a curse of dimensionality associated to their large state space. Finally, in applications and empirical work, social interactions are typically identified, even with population data, only under heroic assumptions.

In this paper we have attempted to show how some of these obstacles to the study of economies with social interactions can be overcome. Admittedly, we have restricted our analysis to linear economies, but in this context we have been able i) to obtain several desirable theoretical properties, like existence, uniqueness, ergodicity; ii) to develop simple recursive methods to rapidly compute equilibria; and iii) to characterize several general properties of dynamic equilibria. Furthermore, while linearity in principle renders the identification problem in static economies with social interaction almost insurmountable, we have been able to exploit the properties of dynamic equilibria in non-stationary economies to produce a positive identification result.

In conclusion, we believe that the class of dynamic linear economies with social interactions we have studied in this paper can be fruitfully and easily employed in applied and empirical work.

9 Bibliography

Aizer, A. and Currie, J. (2004), “Networks or Neighborhood? Correlations in the Use of Publicly-Funded Maternity Care in California,” *Journal of Public Economics*, 88, 2573-2585.

Akerlof, G. (1997), ‘Social Distance and Social Decisions,’ *Econometrica*, 65, 1005-1027.

Akerlof, G. A., Katz, M. L., and Yellen, J. L. (1996), “An Analysis of Out-of-Wedlock Childbearing in the United States,” *Quarterly Journal of Economics*, 111, 277-317.

Ali, M. M. and Dwyer, D. S. (2009), “Estimating Peer Effects in Adolescent Smoking Behavior: A Longitudinal Analysis,” *Journal of Adolescent Health*, 45, 402-408.

Aliprantis, D., and Border, K. C. (2006), *Infinite Dimensional Analysis*, third edition. Springer-Verlag, Berlin.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005), “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools,” *Journal of Political Economy*, 113(1), 151-183.

Ammermueller, A. and Pischke J.S., (2009), “Peer Effects in European Primary Schools: Evidence from PIRLS,” *Journal of Labor Economics*, 27, 3, 315-348.

Axtell, R., S. Durlauf, J.M. Epstein, R. Hammond, B. Klemens, J. Parker, Z. Song, T. Valente, and H.P. Young (2006): “Social Influences and Smoking Behavior: Final Report to the American Legacy Foundation,” mimeo, Center on Social and Economic Dynamics, The Brookings Institution.

Bauman, K. and Ennett, S. (1996), “On the Importance of Peer Influence for Adolescent Drug Use: Commonly Neglected Considerations,” *Addiction*, 91(2):185-198.

Bauwens, L., Laurent, S., and Rombouts, J. V. K. (2006), “Multivariate GARCH models: A Survey,” *Journal of Applied Econometrics*, 21, 1, 79-109.

- Bayer, P., Ross, S. and Topa, G. (2008), "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, Vol. 116 (6), 1150-1196.
- Bearman, P. S., Moody, J., and Stovel, K. (2004), "Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks," *American Journal of Sociology*, 110(1), 44-91.
- Becker, G.S., Grossman, M., and Murphy, K. M., (1994), "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 84, 3, 396-418.
- Becker, G. S. and Murphy, K. M. (1988), "A Theory of Rational Addiction," *Journal of Political Economy*, 96, 4, 675-700.
- Bénabou, R. (1993), "Workings of a City: Location, Education and Production," *Quarterly Journal of Economics*, 108, 619-52.
- Bénabou, R. (1996), "Equity and Efficiency in Human Capital Investment: The Local Connection," *Review of Economic Studies*, 62, 237-264.
- Benhabib, J., Bisin, A., and Jackson, M. O. (eds.) (2011), *Handbook of Social Economics*, Vol. 1A, The Netherlands: North-Holland, 339-416.
- Benveniste, L. M. and Scheinkman, J. A., (1979), "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica*, 47, 727-32.
- Berge, C. (1963), *Topological Spaces*. New York, Macmillan.
- Bernheim, B. D. (1994), 'A Theory of Conformity,' *Journal of Political Economy*, 102, 841-877.
- Bifulco, R., Fletcher, J. and S. Ross (2009), "The Effect of Classmate Characteristics on Individual Outcomes: Evidence from the Add Health", forthcoming in *American Economic Journal: Economic Policy*.
- Binder, M. and Pesaran, M. H. (2001), "Life-Cycle Consumption under Social Interactions," *Journal of Economic Dynamics and Control*, 25, 35-83.
- Bisin, A., Horst, U. and Özgür, O. (2006), "Rational Expectations Equilibria of Economies with Local Interactions," *Journal of Economic Theory*, 127, 74-116.

- Bisin, A., A. Moro, and G. Topa, (2010), "The Empirical Content of Models with Multiple Equilibria in Economies with Social Interactions," mimeo, New York University.
- Bjorn, P.A. and Q.H. Vuong (1985): " Simultaneous Equation Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation," Caltech, Working Paper 537.
- Blume, L. (1982), "New Techniques for the Study of Stochastic Equilibrium Processes," *Journal of Mathematical Economics*, 9, 61-70.
- Blume, L. (2003), "Stigma and Social Control," mimeo Cornell University.
- Blume, L., Brock, W., Durlauf, S., and Ioannides, Y. (2011), "Identification of Social Interactions," in Benhabib, J., Bisin, A., and Jackson, M. O. (eds.): *Handbook of Social Economics*, Vol. 1B, The Netherlands: North-Holland, 855-966.
- Blume, L. and S. Durlauf, (2001), "The Interactions-Based Approach to Socioeconomic Behavior," in S. Durlauf and P. Young (Eds.), *Social Dynamics*, Brookings Institution Press and MIT Press, Cambridge, MA.
- Blume, L. and Durlauf, S. (2005), "Identifying Social Interactions: A Review," in *Methods in Social Epidemiology*, J. M. Oakes and J. Kaufman, eds., San Francisco: Jossey-Bass.
- Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.
- Bollerslev T, Chou, R. Y., and Kroner, K. F. (1992), "ARCH modelling in finance: a review of the theory and empirical evidence," *Journal of Econometrics*, 52, 5-59.
- Bollerslev, T. and Engle, R. F. (1993), "Common Persistence in Conditional Variances," *Econometrica*, 61, 1, 167-186.
- Bollerslev, T, Engle, R. F., Nelson, D. B. (1994), "ARCH models," In *Handbook of Econometrics*, Engle, R. F. and McFadden, D. (eds), North Holland Press: Amsterdam.
- Boozer, M. and Cacciola, S. (2001), "Inside the Black Box of Project Star: Estimation of Peer

Effects Using Experimental Data,” Yale University Economic Growth Center Discussion Paper no. 832.

Borjas, G. (1995), “Ethnicity, neighbourhoods, and human capital externalities,” *American Economic Review*, 85, 365-390.

Boyer, M. (1978), “A Habit Forming Optimal Growth Model,” *International Economic Review*, 19, 3, 585-609.

Boyer, M. (1983), “Rational Demand and Expenditures Patterns under Habit Formation,” *Journal of Economic Theory*, 31, 1, 27-53.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009), “Identification of Peer Effects through Social Networks,” *Journal of Econometrics*, 150, 41-55.

Broadhead, R.S., et al. (1998): “Harnessing peer networks as an instrument for AIDS prevention: Results from a peer-driven intervention,” *Public Health Reports*, 113(Suppl. 1): 42-57.

Brock, W. and Durlauf, S. (2001a), “Discrete Choice with Social Interactions”, *The Review of Economic Studies*, 68, 235-260.

Brock, W. and Durlauf, S. (2001b), “Interactions-Based Models,” in J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Vol. V, North-Holland, Amsterdam.

Brock, W. and Durlauf, S. (2007), “Identification of Binary Choice Models with Social Interactions,” *Journal of Econometrics*, 140, 1, 52-75.

Brock, W. and Durlauf, S. (2010), “Adoption Curves and Social Interactions,” *Journal of the European Economic Association*, 8, 1, 232-251.

Burke, M. and Heiland, F. (2007), “Social Dynamics of Obesity”, *Economic Inquiry* 45, 3, 571-591.

Cabral, L. (1990): “On the Adoption of Innovations with ‘Network’ Externalities,” *Mathematical Social Sciences*, 19, 299-308.

Calvo-Armengol, T. and Jackson, M. (2004), “The Effects of Social Networks on Employment and Inequality,” *American Economic Review*, 94, 426-454.

- Calvo-Armengol, T., E. Patacchini and Y. Zenou (2009), "Peer Effects and Social Networks in Education", *Review of Economic Studies*, 76, 4, 1239-1267.
- Carrell, S. E., Fullerton, R. L., and West, J. E. (2009), "Does Your Cohort Matter? Measuring Peer Effects in College Achievement," *Journal of Labor Economics*, 27 (3), 439-464.
- Case, A., and L. Katz (1991), "The Company You Keep: The Effect of Family and Neighborhood on Disadvantaged Families," *NBER Working Paper* 3705.
- Chaloupka, F. J. and Warner, K. E. (2000), "The economics of smoking," in: A. J. Culyer and J. P. Newhouse (ed.), *Handbook of Health Economics*, vol. 1, chapter 29, 1539-1627, Elsevier.
- Chiappori, P. A. and Ekeland, I. (2009), "The Micro Economics of Efficient Group Behavior: Identification," *Econometrica*, 77, 763-99.
- Christakis, N.A. and J.H. Fowler (2007), "The Spread of Obesity in a Large Social Network over 32 Years", *New England Journal of Medicine*, 357, 370-379.
- Cipollone, P. and Rosolia, A. (2007), "Social Interactions in High School: Lessons from an Earthquake," *American Economic Review*, 97(3), 948-965.
- Clark, A. E. and Lohéac, Y. (2007), "'It wasn't me, it was them!' Social Influence in Risky Behavior by Adolescents," *Journal of Health Economics*, 26, 763-784.
- Cole, H. L., Mailath, G. J. and Postlewaite, A. (1992), "Social Norms, Savings Behaviour, and Growth," *Journal of Political Economy*, 100, 1092-1125.
- Conley, T. and Topa, G. (2002), "Socio-economic Distance and Spatial Patterns in Unemployment," *Journal of Applied Econometrics*, 17, 4, 303-327.
- Conley, T. and Topa, G. (2003), "Identification of Local Interaction Models with Imperfect Location Data," *Journal of Applied Econometrics*, 18, 5, 605-618.
- Conley, T. and Topa, G. (2007), "Estimating Dynamic Local Interactions Models," *Journal of Econometrics*, Vol. 140 (1), 282-303.

- Conti, G., Galeotti, A., Mueller, G. and S. Pudney (2009), "Popularity", ISER Working Paper 2009-03.
- Cook, P. and Moore, M. (2000), "Alcohol," In: Culyer, A.J., Newhouse, J.P. (Eds.), *Handbook of Health Economics*, vol. 1b, Elsevier Science, Netherlands, 1629-1673.
- Crane, J. (1991), "The Epidemic Theory of Ghettoes and Neighborhood Effects of Dropping Out and Teenage Childbearing," *American Journal of Sociology*, 96, 1226-1259.
- Currarini, S., M.O. Jackson, and P. Pin (2009), "An Economic Model of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, 77(4), 1003-45.
- Cutler, D. M. and Glaeser, E. L. (2007), "Social Interactions and Smoking," mimeo Harvard University and NBER working paper No. 13477.
- Davezies, L., d'Haultfoeuille, X., and Fougère, D., (2009), "Identification of Peer Effects Using Group Size Variation," *The Econometrics Journal*, 12, 3, 397-413.
- DeCicca, P., Kenkel, D. S., and Mathios, A. D. (2008), "Cigarette Taxes and the Transition from Youth to Adult Smoking: Smoking Initiation, Cessation, and participation," *Journal of Health Economics*, 27, 4, 904-917.
- De Giorgi, G., Pellizzari, M., and Redaelli, S. (2009), "Be as Careful of the Company You Keep as of the Books You Read: Peer Effects in Education and on the Labor Market," NBER Working Paper No. 14948.
- De Paula, A. (2009), "Inference in a Synchronization Game with Social Interactions," *Journal of Econometrics*, 148, 5671.
- Doob, J. L. (1953), *Stochastic Processes*, New York: Wiley.
- Duffie, D., Geanakoplos, J., Mas-Colell, A., and McLennan, A. (1994), "Stationary Markov Equilibria," *Econometrica*, 62, 745-781.
- Dufour, J-M. and Hsiao, C. (2008), "Identification," *The New Palgrave Dictionary of Economics*, Second Edition, Eds. Steven N. Durlauf and Lawrence E. Blume.

Durlauf, S. (1996), "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," in Barnett, W., G. Gandolfo, and C. Hillinger (Eds.), *Dynamic Disequilibrium Modelling: Proceedings of the Ninth International Symposium on Economic Theory and Econometrics*, Cambridge University Press, New York.

Durlauf, S. N. (2004), "Neighborhood Effects," in *Handbook of Urban and Regional Economics*, vol. 4, ed. J. V. Henderson and J. F. Thisse. Amsterdam: North Holland.

Elster, J. (1999), *Addiction: Entries and Exits*, edited by Jon Elster, New York: Russell Sage Foundation.

Elster, J. and Skog, O. (1999), *Getting Hooked: Rationality and the Addictions*, edited by Jon Elster and Ole-Jørgen Skog, Cambridge: Cambridge University Press.

Engle, R. F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation," *Econometrica*, 50, 987-1008.

Engle, R. F. and Bollerslev, T. (1986), "Modelling the Persistence of Conditional Variances," *Econometric Reviews*, 5, 1-50; 81-87.

Evans, W., W. Oates, and R. Schwab (1992), "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy*, 100, 51, 966-991.

Fletcher, J. M. (2009), "Social Interactions and Smoking: Evidence Using Multiple Student Cohorts, Instrumental Variables, and School Fixed Effects," forthcoming in *Health Economics*.

Föllmer, H. and U. Horst (2001), "Convergence of Locally and Globally Interacting Markov Chains," *Stochastic Processes and Their Applications*, 96, 99-121.

Frisch, R. (1928), "Correlation and scatter in statistical variables," *Nordic Stat. Jour.*, Vol. 1, 36.

Frisch, R. (1933), "Pitfalls in the statistical construction of demand and supply curves," Veröffentlichungen der Frankfurter Ges. für Konjunkturforschung, Neue Folge, Heft 5, Leipzig.

Frisch, R. (1934), *Statistical Confluence Analysis by Means of Complete Regression Systems*, Publ. No. 5, Universitetets Økonomiske Institutt, Oslo.

Frisch, R. (1938), *Statistical Versus Theoretical Relations in Economic Macrodynamics*, mimeographed document for League of Nations conference.

Frisch, R. and Mudgett, B. D. (1931) "Statistical Correlation and the Theory of Cluster Types," *Journal of the American Statistical Association*, Vol. 26, 375.

Futia, C. A. (1982), "Invariant Distributions and the Limiting Behavior of Markovian Economic Models," *Econometrica*, 50, 377-408.

Gaviria, A., and Raphael, S. (2001), "School-based Peer Effects and Juvenile Behavior," *Review of Economics and Statistics*, 83(2), 257-268.

Gilleskie, D. and Strumpf, K. (2005), "The Behavioral Dynamics of Youth Smoking," *Journal of Human Resources*, 40(4), 822-866.

Gilleskie, D. B. and Zhang, Y. S. (2010), "Friendship Formation and Smoking Initiation Among Teens," mimeo UNC and UMichigan.

Glaeser, E., B. Sacerdote, and J. Scheinkman (2003), "The Social Multiplier," *Journal of the European Economic Association*, 1, 345-353.

Glaeser, E., B. Sacerdote, and J. Scheinkman (1996), "Crime and Social Interactions," *Quarterly Journal of Economics*, CXI, 507-548.

Glaeser, E. and J. Scheinkman (2003), "Non-Market Interactions," in *Advances in Economics and Econometrics: Theory and Applications, Eight World Congress, Vol. I*, (M. Dewatripont, L.P. Hansen, and S. Turnovsky, eds.), Cambridge University Press, 339-369.

Glaeser, E. and J. Scheinkman (2001), "Measuring Social Interactions," in S. Durlauf and P. Young (Eds.), *Social Dynamics*, Brookings Institution Press and MIT Press, Cambridge, MA.

Goering, J. and Feins, J. (1997), "The Moving to Opportunity Social Experiment: Early Stages of Implementation and Research Plans," *Poverty Research News*, 1, 4-6.

Goldin, C. and Katz, L. F. (2002), "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions," *Journal of Political Economy*, 110, 4, 730-770.

Graham, B. S., and Hahn, J. (2005), "Identification and Estimation of the Linear-in-Means Model of Social Interactions," *Economics Letters*, 88,1,1-6.

Graham, B. S. (2008), "Identifying Social Interactions through Excess Variance Contrasts," *Econometrica*, 76, 643-660.

Graham, B. (2011), "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complimentarity and Social Spillovers," in In Jess Benhabib, Matthew O. Jackson and Alberto Bisin editors: *Handbook of Social Economics*, Vol. 1B, The Netherlands: North-Holland.

Granovetter, M. (1973), "The Strength of Weak Ties," *American Journal of Sociology*, 78(6), 1360-1380.

Granovetter, M. (1995), *Getting a Job: A Study in Contacts and Careers*, University of Chicago Press.

Grodner, A. Kniesner, T. J. (2007), "Social Interactions in Labor Supply," *Journal of the European Economic Association*, 4:6, 1226-1248

Gruber, J. and Koszegi, B. (2001), "Is addiction "rational"? Theory and evidence," *Quarterly Journal of Economics*, 116, 4, 1261-1303.

Gul, F. and Pesendorfer, W. (2007), "Harmful Addiction," *Review of Economic Studies*, 74, 147-172.

Haavelmo, T. (1944), "The Probability Approach in Econometrics," *Econometrica*, Vol. 12, July.

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, NJ.

Hansen, L. (1982), "Large sample Properties of Generalized Methods of Moment Estimators," *Econometrica*, 50, 1029-1056.

Hausman, J. A. and Taylor, W. E. (1983), "Identification in Linear Simultaneous Equations Models with Covariance Restrictions: An Instrumental Variables Interpretation," *Econometrica*, 51 (5), 1527-1549.

Hansen, L. and Sargent, T. (2004), *Recursive Models of Dynamic Linear Economies*, University

of Chicago, mimeo.

Hoff, K. and Sen, A. (2005), "Homeownership, Community Interactions, and Segregation," *American Economic Review*, 95, 1167-1189.

Horst, U. and Scheinkman, J. (2006), "Equilibria in Systems of Social Interactions," *Journal of Economic Theory*, 130, 44-77.

Hoxby, C. M. (2000a), "The Effects Of Class Size On Student Achievement: New Evidence From Population Variation," *The Quarterly Journal of Economics*, 115, 4, 1239-1285.

Hoxby, C. M. (2000b), "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER Working Paper 7867.

Hsiao, C. (1983), "Identification," In *Handbook of Econometrics*, vol. 1, ed. Z. Griliches and M.D. Intriligator. Amsterdam: North-Holland.

Hurwicz, L. (1950), "Generalization of the Concept of Identification," *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph 10, New York: John Wiley and Sons.

Iannaccone, L. R. (1986), "Addiction and Satiation," *Economics Letters*, 21, 1, 95-99.

Ichino, A. and Falk, A. (2006), "Clean Evidence on Peer Pressure," *Journal of Labor Economics*, 24, 39-57.

Ioannides, Y. and L. Datcher Loury (2004), "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 42, 1056-93.

Ioannides, Y. and Topa, G. (2009), "Neighborhood Effects: Accomplishments and Looking Beyond Them," *forthcoming in Journal of Regional Science*.

Ioannides, Y., and J. Zabel (2008), "Interactions, Neighborhood Selection, and Housing Demand," *Journal of Urban Economics*, 63, 229-252.

Ioannides, Y. and Zanella, G. (2009), "Neighborhood Effects and Neighborhood Choice," mimeo.

Jones, S. R. G. (1984), *The Economics of Conformism*, Oxford: Basil Blackwell.

Kallenberg, O. (2002), *Foundations of Modern Probability*, second edition. Springer-Verlag, Berlin.

Kandel, D. B. (1978), "Homophily, Selection, and Socialization in Adolescent Friendships," *American Journal of Sociology*, 84, 427-436.

Katz, L., Kling, A. and Liebman, J. (2001), "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics*, CXVI, 607-654.

Kling, J., Ludwig, J. and Katz, L. (2005), "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment," *Quarterly Journal of Economics*, 120(1), 87-130.

Kobus, K. (2003), "Peers and adolescent smoking," *Addiction*, 98, Supplement 1, 37-55.

Koopmans, T. C. (1949), "Identification Problems in Economic Model Construction," *Econometrica*, 17, 125-144.

Koopmans, T. C., Rubin, H., and Leipnik, R. B. (1950), "Measuring the Equation Systems of Dynamic Economics," *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph 10, New York: John Wiley and Sons.

Koopmans, T. C. and Reiersøl, O. (1950), "The Identification of Structural Characteristics," *Annals of Mathematical Statistics*, 21, 165-181.

Krauth, B. V. (2005), "Peer effects and selection effects on smoking among Canadian youth," *Canadian Journal of Economics*, 38, 735-757.

Krauth, B. V. (2006), "Simulation-based estimation of peer effects," *Journal of Econometrics*, 133, 243-271.

Kremer, M. and Levy, D. (2008), "Peer Effects and Alcohol Use among College Students," *The Journal of Economic Perspectives*, 22,3, 189-206.

Krosnick, J. and C. Judd (1982), "Transitions in Social Influence in Adolescence: Who Induces Cigarette Smoking," *Developmental Psychology*, 81, 359-68.

- Leider, S., Mobius, M. M., Rosenblat, T., and Do, Q-A. (2007), "How Much is A Friend Worth? Directed Altruism and Enforced Reciprocity in Social Networks," Boston Fed working paper.
- Lewitt, E., Coate, D. and Grossman, M. (1981), "The Effects of Government Regulation on Teenage Smoking," *The Journal of Law and Economics*, 24, 545-569.
- Liggett, T.M. (1985) *Interacting Particle Systems*, Springer Verlag, Berlin.
- Ludwig, J., Hirschfeld, P. and Duncan, G. (2001), "Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment," *Quarterly Journal of Economics*, CXVI (2): 665-679.
- Lundborg, P. (2006), "Having the wrong friends? Peer effects in adolescent substance use," *Journal of Health Economics*, 25, 214-233.
- Mailath, G. J. and Samuelson, L. (2006), *Repeated Games and Reputations: Long-run Relationships*, Oxford University Press, New York.
- Mandelbrot, B. (1967), "Some Noises with $1/f$ Spectrum, a Bridge between Direct Current and White Noise," *IEEE Transactions on Information Theory*, vol. IT-13, 2, 289-298.
- Manski, C. (1993), "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531-42.
- Manski, C. (2000), "Economic Analysis of Social Interactions," *Journal of Economic Perspectives*, 14, 115-136.
- Manski, C. (2007), *Identification for Prediction and Decision*, Harvard University Press, Cambridge, MA.
- Marschak, J. (1942), "Economic Interdependence and Statistical Analysis," *Studies in Mathematical Economics and Econometrics*, University of Chicago Press, 135-150.
- Matzkin, R. (2007), "Nonparametric Identification," In *Handbook of Econometrics*, vol. 6, ed. J. Heckman and E. Leamer. Amsterdam: North-Holland.

- Mihaly, K., (2007) "Too Popular for School? Friendship Formation and Academic Achievement", unpublished manuscript, Duke University.
- Mobius, M. (2000), "The Formation of Ghettos as a Local Interaction Phenomenon," mimeo, Harvard University.
- Moffitt, R. A. (2001), 'Policy Interventions, Low-Level equilibria, and Social Interactions,' in S. Durlauf and P. Young (Eds.), *Social Dynamics*, Brookings Institution Press and MIT Press, Cambridge, MA.
- Moody, J., (2001), "Race, School Integration, and Friendship Segregation in America", *American Journal of Sociology*, 107, 3, 679–716.
- Moro, A. (2003), "The Effect of Statistical Discrimination on Black-White Inequality: Estimating a Model with Multiple Equilibria", *International Economic Review*, 44, 2, 467-500.
- Nakajima, R. (2007), "Measuring Peer Effects on Youth Smoking Behavior", *Review of Economic Studies*, 74, 897-935.
- Neveu, J. (1965), *Mathematical Foundations of the Calculus of Probability*, Translated by Amiel Feinstein, San Francisco: Holden-Day.
- Norton, E.C., R.C. Lindrooth, and S.T. Ennett (1998): "Controlling for the Endogeneity of Peer Substance Use on Adolescent Alcohol and Tobacco Use," *Health Economics*, 7(5), 439-53.
- Peski, M. (2007), "Complementarities, Group Formation, and Preferences for Similarity," mimeo.
- Pigou, A. C. (1910), "A Method of Determining the Numerical Values of Elasticities of Demand," *Economic Journal*, Vol. 20, 636-640.
- Powell, L.M., J.A. Tauras, and H. Ross (2003): "Peer Effects, Tobacco Control Policies, and Youth Smoking Behavior," University of Illinois at Chicago, mimeo.
- Rosenbaum, J. (1995), "Changing the geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program," *Housing Policy Debate*, 6, 231-269.
- Rothenberg, T. J. (1971), "Identification in Parametric Models," *Econometrica*, 39 (3), 577-591.

- Rozen, K. (2010), "Foundations of Intrinsic Habit Formation," *Econometrica*, 78 (4), 1341-1373.
- Sacerdote, B. (2001), "Peer Effects with Random Assignments: Results from Dartmouth Roommates," *Quarterly Journal of Economics*, 116, 681-704.
- Santos, M. S. (1991), "Smoothness of the Policy Function in Discrete Time Economic Models," *Econometrica*, 59, 1365-82.
- Schelling, T. (1971), "Dynamic Models of Segregation," *Journal of Mathematical Sociology*, 1, 143-186.
- Schelling, T. (1972), "A Process of Residential Segregation: Neighborhood Tipping," in *Racial Discrimination in Economic Life*, ed. by A. Pascal, Lexington Books, Lexington, MA.
- Schultz, H. (1938), *Theory and Measurement of Demand*, University of Chicago Press.
- Soetevent, A. and Kooreman, P. (2007), "A Discrete Choice Model with Social Interactions; With an Application to High School Teen Behavior," *Journal of Applied Econometrics*, vol. 22, issue 3, 599-624.
- Stokey, N., Lucas, R. E. with Prescott, E. (1989): *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, MA.
- Topa, G. (2001), "Social Interactions, Local Spillover and Unemployment," *The Review of Economic Studies*, 68, 261-95.
- Tweedie, R. L. (1975), "Sufficient Conditions for Ergodicity and Recurrence of Markov Chains in a General State Space," *Stochastic Processes and Their Applications*, 3, 385-403.
- Tyas, S.L. and L.L. Pederson (1998): "Psychosocial Factors Related to Adolescent Smoking: A Critical Review of the Literature," *Tobacco Control*, 7(4), 409-20.
- Wald, A. (1950), "Note on the Identification of Economic Relations," *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph 10, New York: John Wiley and Sons.
- Wang, M.Q., E.C. Fitzhugh, R.C. Westerfield, and J.M. Eddy (1995): "Family and Peer Influ-

ences on Smoking Behavior among American Adolescent: An Age Trend,” *Journal of Adolescent Health*, 16(3), 200-03.

Wang, M.Q., J.M. Eddy, E.C. Fitzhugh, (2000): ”Smoking Acquisition: Peer Influence and Self-selection,” *Psychological Reports*, 86, 1241-46.

Weinberg, B., Reagan, P. and Yankow, J. (2004), “Do Neighborhoods Affect Hours Worked? Evidence from Longitudinal Data,” *Journal of Labor Economics*, 22, 4, 891-924.

Young, P. (2009), “Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning,” *American Economic Review*, 99, 1899-1924.

Zanella, G. (2007), “Discrete Choice with Social Interactions and Endogenous Membership,” *Journal of the European Economic Association*, 5, 122-153.

Zimmerman, D. (2003), “Peer Effects in Higher Education: Evidence From a Natural Experiment,” *Review of Economics and Statistics*, 85, 9-23.

10 Appendix A: The existence proof (Theorem 1)

The proof is constructive and works in three steps, by induction on the length of the continuation economies.⁶¹

Step 1: Existence, uniqueness and the Markov property for $T = 1$. In this symmetric environment, it is enough to analyze the optimization problem of a single agent, say of agent $0 \in \mathbb{A}$. We will allow for arbitrary initial histories so that one can interpret the current step either as a one-period economy or the last period of a finite-horizon economy. We will show that, agents will use only the information contained in the previous period choices x_0 and current type realizations θ_1 . Let any t -length history $(x^{t-1}, \theta^t) = (x_{-(t-1)}, \theta_{-(t-2)}, \dots, x_{-1}, \theta_0, x_0, \theta_1)$ of previous choices and preference shock realizations be given. Agent 0 solves

$$\max_{x_1^0 \in X} \left\{ -\alpha_1 (x_0^0 - x_1^0)^2 - \alpha_2 (\theta_1^0 - x_1^0)^2 - \alpha_3 (x_1^{-1} - x_1^0)^2 - \alpha_3 (x_1^1 - x_1^0)^2 \right\} \quad (23)$$

The first order condition

$$2 [\alpha_1 (x_0^0 - x_1^0) + \alpha_2 (\theta_1^0 - x_1^0) + \alpha_3 (x_1^{-1} - x_1^0) + \alpha_3 (x_1^1 - x_1^0)] = 0$$

implies that

$$x_1^0 = \Delta_1^{-1} (\alpha_1 x_0^0 + \alpha_2 \theta_1^0 + \alpha_3 x_1^{-1} + \alpha_3 x_1^1) \quad (24)$$

where

$$\Delta_1 := (\alpha_1 + \alpha_2 + 2\alpha_3) > 0$$

This choice is feasible (in X) since it is a convex combination of elements of X , a convex set by assumption. The objective function (23) is strictly concave in x_1^0 , thus x_1^0 in (24) is the unique optimizer. We see from (24) that showing the existence of a symmetric equilibrium in the continuation given history (x^{t-1}, θ^t) is equivalent to finding the fixed point of an operator $L_1 : B((\mathbf{X} \times \Theta)^t, X) \rightarrow B((\mathbf{X} \times \Theta)^t, X)$ that acts on the class of bounded measurable functions $x_1 : (\mathbf{X} \times \Theta)^t \rightarrow X$ according to

$$(L_1 x_1)(x^{t-1}, \theta^t) = \Delta_1^{-1} (\alpha_1 x_0^0 + \alpha_2 \theta_1^0 + \alpha_3 x_1(R^{-1} x^{t-1}, R^{-1} \theta^t) + \alpha_3 x_1(R x^{t-1}, R \theta^t))$$

Clearly, L_1 is a self-map. We show next that it is a contraction. Endow $B((\mathbf{X} \times \Theta)^t, X)$ with the sup norm which makes $(B((\mathbf{X} \times \Theta)^t, X), \|\cdot\|_\infty)$ a Banach space. Pick $x_1, \hat{x}_1 \in B((\mathbf{X} \times \Theta)^t, X)$. We have

⁶¹We laid out the problem in its recursive form for clarity in Section 2. Our method of proof attacks the sequence problem directly.

for all (x^{t-1}, θ^t)

$$\begin{aligned}
\left| (L_1 x_1)(x^{t-1}, \theta^t) - (L_1 \hat{x}_1)(x^{t-1}, \theta^t) \right| &= \Delta_1^{-1} \left| \alpha_1 x_0^0 + \alpha_2 \theta_1^0 \right. \\
&\quad + \alpha_3 x_1(R^{-1} x^{t-1}, R^{-1} \theta^t) + \alpha_3 x_1(R x^{t-1}, R \theta^t) \\
&\quad - \alpha_1 x_0^0 - \alpha_2 \theta_1^0 \\
&\quad \left. - \alpha_3 \hat{x}_1(R^{-1} x^{t-1}, R^{-1} \theta^t) - \alpha_3 \hat{x}_1(R x^{t-1}, R \theta^t) \right| \\
&= \Delta_1^{-1} \left| \alpha_3 (x_1(R^{-1} x^{t-1}, R^{-1} \theta^t) - \hat{x}_1(R^{-1} x^{t-1}, R^{-1} \theta^t)) \right. \\
&\quad \left. + \alpha_3 (x_1(R x^{t-1}, R \theta^t) - \hat{x}_1(R x^{t-1}, R \theta^t)) \right| \\
&\leq \left(\frac{\alpha_3}{\Delta_1} \right) \left| x_1(R^{-1} x^{t-1}, R^{-1} \theta^t) - \hat{x}_1(R^{-1} x^{t-1}, R^{-1} \theta^t) \right| \\
&\quad + \left(\frac{\alpha_3}{\Delta_1} \right) \left| x_1(R x^{t-1}, R \theta^t) - \hat{x}_1(R x^{t-1}, R \theta^t) \right| \\
&\leq \left(\frac{2\alpha_3}{\Delta_1} \right) \|x_1 - \hat{x}_1\|_\infty
\end{aligned}$$

The coefficient $2\alpha_3 \Delta_1^{-1} < 1$ since $\alpha_i > 0$, for $i = 1, 2, 3$. Hence L_1 is a contraction on $B((\mathbf{X} \times \Theta)^t, X)$. Thus, by Banach Fixed Point Theorem (see e.g., Aliprantis and Border (2006), p.95) L_1 has a unique fixed point x_1^* in $B((\mathbf{X} \times \Theta)^t, X)$. Next, we argue that this equilibrium strategy must be Markovian and should assume the convex combination form as in the statement of Theorem 1.

Lemma 1 (Markov Property and the Convex Combination Form) *Unique symmetric equilibrium strategy x_1^* is Markovian, i.e., it depends solely on last period equilibrium choices and current preference shock realizations: for any t -length history (x^{t-1}, θ^t) , $x_1^*(x^{t-1}, \theta^t) = g_1(x_0, \theta_1)$, for some $g_1 : \mathbf{X} \times \Theta \rightarrow X$. Moreover, the Markovian policy function g_1 has the convex combination form as in the statement of the theorem.*

Proof: Let

$$G := \left\{ \begin{array}{l} g : \mathbf{X} \times \Theta \rightarrow X \text{ s.t. } g(x, \theta) = \sum_{a \in \mathbb{A}} c^a x^a + \sum_{a \in \mathbb{A}} d^a \theta^a + e \bar{\theta} \\ \text{with} \\ \text{(i) } c^a, d^a, e \geq 0 \text{ and } e + \sum_{a \in \mathbb{A}} (c^a + d^a) = 1 \\ \text{(ii) } (\frac{1}{2})c^{a+1} + (\frac{1}{2})c^{a-1} \geq c^a, \forall a \neq 0 \\ \text{(iii) } c^b \leq c^a, \forall a, b \in \mathbb{A} \text{ with } |b| > |a|. \\ \text{(iv) } c^a = c^{-a}, \forall a \in \mathbb{A} \\ \text{and properties (ii), (iii), and (iv) also holding for the } d = (d^a)_{a \in \mathbb{A}} \text{ sequence.} \end{array} \right\} \quad (25)$$

be the class of functions that are convex combinations (i) of one-period before history, current types and average type, having the (ii) ‘convexity’, (iii) ‘monotonicity’, and (iv) ‘symmetry’ properties. Property (ii) states that the rate of ‘spatial’ (cross-sectional) convergence of the policy weights is non-increasing in both directions, from the center. Monotonicity property, (iii), has a very natural interpretation: agent b ’s effect on agent 0’s marginal utility is smaller than agent a ’s effect on it, if a is closer to 0 than b is. Finally, (iv) says that the policy weights are symmetric around 0. Let $g \in G$ be such that after any history (x^{t-1}, θ^t)

$$x_1(x^{t-1}, \theta^t) = g(x_0, \theta_1)$$

and let (c, d, e) be the coefficient sequence associated with g . Applying L_1 to x_1 (hence to g), we get

$$\begin{aligned}
(L_1 x_1)(x^{t-1}, \theta^t) &= \Delta_1^{-1} (\alpha_1 x_0^0 + \alpha_2 \theta_1^0 + \alpha_3 g(R^{-1} x_0, R^{-1} \theta_1) + \alpha_3 g(R x_0, R \theta_1)) \\
&= \Delta_1^{-1} \left(\alpha_1 x_0^0 + \alpha_2 \theta_1^0 \right. \\
&\quad \left. + \alpha_3 \left(\sum_{a \in \mathbb{A}} c^a x_0^{a-1} + \sum_{a \in \mathbb{A}} d^a \theta_1^{a-1} + e \bar{\theta} \right) + \alpha_3 \left(\sum_{a \in \mathbb{A}} c^a x_0^{a+1} + \sum_{a \in \mathbb{A}} d^a \theta_1^{a+1} + e \bar{\theta} \right) \right)
\end{aligned} \tag{26}$$

Reorganizing the terms gives

$$\begin{aligned}
&= \Delta_1^{-1} \left(\alpha_1 x_0^0 + \alpha_2 \theta_1^0 \right. \\
&\quad \left. + \alpha_3 \left(\sum_{a \in \mathbb{A}} c^{a+1} x_0^a + \sum_{a \in \mathbb{A}} d^{a+1} \theta_1^a + e \bar{\theta} \right) + \alpha_3 \left(\sum_{a \in \mathbb{A}} c^{a-1} x_0^a + \sum_{a \in \mathbb{A}} d^{a-1} \theta_1^a + e \bar{\theta} \right) \right)
\end{aligned}$$

and rearranging gives

$$\begin{aligned}
&= \Delta_1^{-1} \left(x_0^0 \underbrace{(\alpha_1 + \alpha_3 c^{-1} + \alpha_3 c^1)}_{\Delta_1 \hat{c}^0} + \theta_1^0 \underbrace{(\alpha_2 + \alpha_3 d^{-1} + \alpha_3 d^1)}_{\Delta_1 \hat{d}^0} + 2\alpha_3 e \bar{\theta} \right. \\
&\quad \left. + \sum_{a \neq 0} \underbrace{(\alpha_3 c^{a-1} + \alpha_3 c^{a+1})}_{\Delta_1 \hat{c}^a} x_0^a + \sum_{a \neq 0} \underbrace{(\alpha_3 d^{a-1} + \alpha_3 d^{a+1})}_{\Delta_1 \hat{d}^a} \theta_1^a \right)
\end{aligned} \tag{27}$$

The function after the last equality sign is linear in x_0, θ_1 and $\bar{\theta}$. So, $L_1 x_1$ preserves the same linear form. By definition of the new coefficient sequence $(\hat{c}, \hat{d}, \hat{e})$ in (27), each element of the sequence is nonnegative since each element of the original one was so. New coefficients sum up to 1 since convex combination form of g makes the sum of the coefficients inside the two parentheses on the right hand side of (26) equal to 1. Thus, the total sum of coefficients on the right hand side of (26) is $\Delta_1^{-1}(\alpha_1 + \alpha_2 + 2\alpha_3) = 1$, which proves property (i). The final form in (27) is just a regrouping of elements in (26). Let $(\hat{c}^a)_{a \in \mathbb{A}}$ be the new coefficient sequence associated with $L_1 x_T$ as defined in equation (27). Pick $a \neq 0$ in $\mathbb{A}$,

$$\begin{aligned}
\hat{c}^{a+1} + \hat{c}^{a-1} &\geq \left(\frac{\alpha_3}{\Delta_1} \right) (c^a + c^{a+2}) + \left(\frac{\alpha_3}{\Delta_1} \right) (c^{a-2} + c^a) \\
&\geq \left(\frac{\alpha_3}{\Delta_1} \right) (2c^{a+1} + 2c^{a-1}) \\
&= 2 \left(\frac{\alpha_3}{\Delta_1} \right) (c^{a+1} + c^{a-1}) \\
&= 2\hat{c}^a
\end{aligned}$$

By definition of $\hat{c}$ in (27), first inequality is strict if $|a| = 1$, is an equality otherwise; second inequality is by property (ii) on c ; last equality is once again by definition of $\hat{c}$ in (27). Therefore, for any $a \neq 0$ in $\mathbb{A}$, $\hat{c}^{a+1} + \hat{c}^{a-1} \geq 2\hat{c}^a$, which is property (ii). Pick any $a, b \in \mathbb{A}$ with $|a| < |b|$.

$$\begin{aligned}
\hat{c}^a &= \left(\frac{\alpha_3}{\Delta_1} \right) c^{a-1} + \left(\frac{\alpha_3}{\Delta_1} \right) c^{a+1} \\
&= \left(\frac{\alpha_3}{\Delta_1} \right) c^{|a|-1} + \left(\frac{\alpha_3}{\Delta_1} \right) c^{|a|+1} \\
&\geq \left(\frac{\alpha_3}{\Delta_1} \right) c^{|b|-1} + \left(\frac{\alpha_3}{\Delta_1} \right) c^{|b|+1} \\
&= \left(\frac{\alpha_3}{\Delta_1} \right) c^{b-1} + \left(\frac{\alpha_3}{\Delta_1} \right) c^{b+1} \\
&= \hat{c}^b
\end{aligned}$$

First equality is from (27); second by property (iv) of G in (25); the inequality is property (iii) of G in (25); next equality is due to property (iv) of G again; and finally the last equality is by (27). Hence, property (iii) in (25) holds for the new sequence. We next show that $\hat{c}$ satisfies (iv) in (25).

$$\begin{aligned}
\hat{c}^a &= \left(\frac{\alpha_3}{\Delta_1} \right) c^{a-1} + \left(\frac{\alpha_3}{\Delta_1} \right) c^{a+1} \\
&= \left(\frac{\alpha_3}{\Delta_1} \right) c^{-a-1} + \left(\frac{\alpha_3}{\Delta_1} \right) c^{-a+1} \\
&= \hat{c}^{-a}
\end{aligned}$$

where first equality is by (27); the second is due to (iv) of G in (25); finally the last is again by (27). Thus, the restriction of L_1 to the space of bounded measurable functions that agree with an element of G after any history (call it B_G), maps elements of this latter into itself. Moreover, endowed with the sup norm, B_G is a closed subset of $B((\mathbf{X} \times \Theta)^t, X)$ since it is defined by equality and inequality constraints, hence a complete metric space in its own right. Since L_1 is a contraction on this latter as we just showed, it is so on B_G too and the unique fixed point x_1^* in $B((\mathbf{X} \times \Theta)^t, X)$ must lie in B_G . Since the choice of t was arbitrary, the unique symmetric equilibrium in a one-period (continuation) economy, after any length history must be Markovian and should assume the convex combination form stated in the theorem ($x_1^*(x^{t-1}, \theta^t) = g_1(x_0, \theta_1)$). This concludes the proof of Lemma 1. $\blacksquare$

This proves Step 1, namely that the statement of the Theorem is true for 1-period continuation economies. Next, we prove that this result generalizes to any finite-horizon economy.

Step 2: Induction, T-1 implies T. Let a T -period finite-horizon economy be given, with $T \geq 2$. Assume that the statement of Theorem 1 is true up to $T-1$ -period. The T -period economy can be *separated* into a first period and a $T-1$ -period continuation economy. Then, by hypothesis, there exists a unique symmetric MPE, $g : \mathbf{X} \times \Theta \times \{1, \dots, T-1\} \mapsto X$, for the $T-1$ -period continuation economy. Agent 0 believes that all other agents, including his own reincarnations, will use that unique symmetric equilibrium map from period 2 on. Given any t -length history (x^{t-1}, θ^t) , agent 0 solves

$$\begin{aligned}
&\max_{x_1^0 \in X} \left\{ -\alpha_1 (x_0^0 - x_1^0)^2 - \alpha_2 (\theta_1^0 - x_1^0)^2 - \alpha_3 (x_1^{-1} - x_1^0)^2 - \alpha_3 (x_1^1 - x_1^0)^2 \right. \\
&\quad \left. + E \left[\sum_{\tau=2}^T \beta^{\tau-1} \left(-\alpha_1 (x_{\tau-1}^0 - x_\tau^0)^2 - \alpha_2 (\theta_\tau^0 - x_\tau^0)^2 - \alpha_3 (x_\tau^{-1} - x_\tau^0)^2 - \alpha_3 (x_\tau^1 - x_\tau^0)^2 \right) \middle| (x^{t-1}, \theta^t) \right] \right\}
\end{aligned} \tag{28}$$

The hypothesis that agents' optimal future choices are given by the policy function g can be used to characterize the forms of the optimal choices on the equilibrium path as composite linear functions as in the following Lemma⁶².

Lemma 2 (Convexity and Monotonicity) *Given a T -period economy, equilibrium choices satisfy the following properties:*

- (i) *For any period $t \geq 2$, agent a 's period t equilibrium choice, x_t^a , can be written as a non-negative weighted sum of the first period choices and the realized paths of the type shocks and their expected value, i.e.,*⁶³

$$\begin{aligned} x_t^a &= \sum_{b_1 \in \mathbb{A}} \cdots \sum_{b_{t-1} \in \mathbb{A}} c_{T-(t-1)}^{b_1} \cdots c_{T-1}^{b_{t-1}} x_1^{a+b_1+\cdots+b_{t-1}} \\ &+ \sum_{s=1}^{t-1} \sum_{b_1 \in \mathbb{A}} \cdots \sum_{b_{s-1} \in \mathbb{A}} c_{T-(t-1)}^{b_1} \cdots c_{T-(t-(s-1))}^{b_{s-1}} \left(\sum_{b_s \in \mathbb{A}} d_{T-(t-s)}^{b_s} \theta_{t-(s-1)}^{a+b_1+\cdots+b_s} + e_{T-(t-s)} \bar{\theta} \right) \end{aligned} \quad (29)$$

- (ii) *The impact on agent 0's period t equilibrium choice of agent a 's first period choice is decreasing with respect to the social distance of agent a to agent 0, i.e., for any $a, b \in \mathbb{A}$, any $t \geq 2$,*

$$|a| \leq |b| \implies \frac{\partial x_t^0}{\partial x_1^a} \leq \frac{\partial x_t^0}{\partial x_1^b}$$

- (iii) *Agent a 's first period choice's average impact on agents -1 and 1 's period t equilibrium choices is greater than his impact on agent 0's, i.e.,*

$$\frac{\partial}{\partial x_1^a} (2x_t^0 - x_t^1 - x_t^{-1}) \leq 0$$

- (iv) *The impact of agent 0's own first period choice on his future choices declines geometrically across periods, i.e.,*

$$\frac{\partial}{\partial x_1^0} x_t^0 \leq \left(\frac{\alpha_1}{\alpha_1 + \alpha_2} \right) \frac{\partial}{\partial x_1^0} x_{t-1}^0$$

Thanks to the linearity of the optimal future choices as shown in Lemma 2, agent 0's problem is differentiable with respect to x_1^0 and the unconstrained ($x_1^0 \in \mathbb{R}$) first order condition for (28) is

$$\begin{aligned} 0 &= \alpha_1 (x_0^0 - x_1^0) + \alpha_2 (\theta_1^0 - x_1^0) + \alpha_3 (x_1^{-1} - x_1^0) + \alpha_3 (x_1^1 - x_1^0) \\ &+ E \left[\sum_{\tau=2}^T \beta^{\tau-1} \left(-\alpha_1 (x_{\tau-1}^0 - x_\tau^0) \frac{\partial}{\partial x_1^0} (x_{\tau-1}^0 - x_\tau^0) + \alpha_2 (\theta_\tau^0 - x_\tau^0) \frac{\partial}{\partial x_1^0} x_\tau^0 \right. \right. \\ &\left. \left. - \alpha_3 (x_\tau^{-1} - x_\tau^0) \frac{\partial}{\partial x_1^0} (x_\tau^{-1} - x_\tau^0) - \alpha_3 (x_\tau^1 - x_\tau^0) \frac{\partial}{\partial x_1^0} (x_\tau^1 - x_\tau^0) \right) \middle| (x^{t-1}, \theta^t) \right] \end{aligned} \quad (30)$$

⁶²Unless otherwise stated, the proofs of the Lemmas are relegated to Appendix D: The Technical Appendix in order to make the reading uninterrupted.

⁶³We use in expression (29) the convention that in the sum after the plus sign, for $s = 1$, the summand becomes $\sum_{b_s \in \mathbb{A}} d_{T-(t-s)}^{b_s} \theta_{t-(s-1)}^{a+b_1+\cdots+b_s} + e_{T-(t-s)} \bar{\theta}$.

Agent 0's problem is strictly concave in his choice x_1^0 since the second partial of the objective function in (28) with respect to x_1^0 , $-\Delta_T$ by definition, is negative, or

$$\begin{aligned} \Delta_T &:= \alpha_1 + \alpha_2 + 2\alpha_3 \\ &+ \sum_{\tau=2}^T \beta^{\tau-t} \left(\alpha_1 \left(\frac{\partial}{\partial x_1^0} (x_{\tau-1}^0 - x_\tau^0) \right)^2 + \alpha_2 \left(\frac{\partial}{\partial x_1^0} x_\tau^0 \right)^2 \right. \\ &\left. + \alpha_3 \left(\frac{\partial}{\partial x_1^0} (x_\tau^{-1} - x_\tau^0) \right)^2 + \alpha_3 \left(\frac{\partial}{\partial x_1^0} (x_\tau^1 - x_\tau^0) \right)^2 \right) > 0 \end{aligned} \quad (31)$$

Consequently, the FOC characterizes the *unique maximizer* of the unconstrained problem ($x_1^0 \in \mathbb{R}$). The following Lemma shows that equation (30) has a much simpler representation.

Lemma 3 (Interiority) *Equation (30) can be written in the following alternative form*

$$0 = -x_1^0 \Delta_T + \alpha_1 x_0^0 + \alpha_2 \theta_1^0 + \sum_{a \neq 0} \gamma_T^a x_1^a + \mu_T \bar{\theta} \quad (32)$$

where $\Delta_T := \alpha_1 + \alpha_2 + \sum_{a \neq 0} \gamma_T^a + \mu_T$. Moreover, the coefficients α_1 , α_2 , $(\gamma_T^a)_{a \neq 0}$, and μ_T are all non-negative.

By isolating the choice x_1^0 on the left hand side, we have

$$x_1^0 = \Delta_T^{-1} \left(\alpha_1 x_0^0 + \alpha_2 \theta_1^0 + \sum_{a \neq 0} \gamma_T^a x_1^a + \mu_T \bar{\theta} \right) \quad (33)$$

which means that the maximizer of the unconstrained problem is a convex combination of x_0^0 , θ_1^0 , $(x_1^a)_{a \neq 0}$ and $\bar{\theta}$. Each of these are elements of X , a convex set. Hence, the maximizer of the unconstrained problem is in the feasible set of the constrained problem. Consequently, it is the unique maximizer of (28). The form in (33) implies that showing the existence of a symmetric equilibrium policy for the first period of a T -period economy is equivalent to finding the fixed point of an operator $L_T : B((\mathbf{X} \times \boldsymbol{\Theta})^t, X) \rightarrow B((\mathbf{X} \times \boldsymbol{\Theta})^t, X)$ that acts on the class of bounded measurable functions $x_1 : (\mathbf{X} \times \boldsymbol{\Theta})^t \rightarrow X$ according to

$$(L_T x_1)(x^{t-1}, \theta^t) = \Delta_T^{-1} \left(\alpha_1 x_0^0 + \alpha_2 \theta_1^0 + \sum_{a \neq 0} \gamma_T^a x_1(R^a x^{t-1}, R^a \theta^t) + \mu_T \bar{\theta} \right)$$

Clearly L_T is a self-map. Using straightforward modifications of the arguments in the proof of **Step 1**, we can show for $x_1, \hat{x}_1 \in B((\mathbf{X} \times \boldsymbol{\Theta})^t, X)$ that

$$\left| (L_T x_1)(x^{t-1}, \theta^t) - (L_T \hat{x}_1)(x^{t-1}, \theta^t) \right| \leq \sum_{a \neq 0} \left(\frac{\gamma_T^a}{\Delta_T} \right) \|x_1 - \hat{x}_1\|_\infty$$

The coefficient $\sum_{a \neq 0} \left(\frac{\gamma_T^a}{\Delta_T} \right) < 1$ since $\alpha_i > 0$, $i = 1, 2, 3$. Thus, L_T is a contraction on the Banach space of bounded measurable functions $(B((\mathbf{X} \times \boldsymbol{\Theta})^t, X), \|\cdot\|_\infty)$, consequently has a unique fixed point x_1^* . Once again, straightforward modifications of the arguments in Lemma 1 yield that perfect equilibria are necessarily Markovian thus we can focus attention on Markovian strategies. As in the proof of Lemma

1, it suffices to show that $L_T(B_G) \subset B_G$. To that effect, let $x_1 \in B_G$ be such that there exists a $g \in G$ such that after any history (x^{t-1}, θ^t) , $x_1(x^{t-1}, \theta^t) = g(x_0, \theta_1)$; let (c, d, e) be the coefficient sequence associated with g . Applying L_T to x_1 , we get

$$\begin{aligned}
(L_T x_1)(x^{t-1}, \theta^t) &= \Delta_T^{-1} \left(\alpha_1 x_0^0 + \alpha_2 \theta_1^0 + \sum_{a \neq 0} \gamma_T^a g(R^a x_0, R^a \theta_1) + \mu_T \bar{\theta} \right) \\
&= \Delta_T^{-1} \left(\alpha_1 x_0^0 + \alpha_2 \theta_1^0 \right. \\
&\quad \left. + \sum_{a \neq 0} \gamma_T^a \left(\sum_{a_1 \in \mathbb{A}} c^{a_1} x_0^{a+a_1} + \sum_{a_2 \in \mathbb{A}} d^{a_2} \theta_1^{a+a_2} + e \bar{\theta} \right) + \mu_T \bar{\theta} \right) \quad (34) \\
&= \Delta_T^{-1} \left(\underbrace{[\alpha_1 + \sum_{a \neq 0} \gamma_T^a c^{-a}] x_0^0}_{\Delta_T \hat{c}^0} + \underbrace{[\alpha_2 + \sum_{a_1 \neq 0} \gamma_T^{a_1} d^{-a_1}] \theta_1^0}_{\Delta_T \hat{d}^0} \right. \\
&\quad \left. + \sum_{b \neq 0} \left\{ \underbrace{[\sum_{a \neq 0} \gamma_T^a c^{b-a}] x_0^b}_{\Delta_T \hat{c}^b} + \underbrace{[\sum_{a \neq 0} \gamma_T^a d^{b-a}] \theta_1^b}_{\Delta_T \hat{d}^b} \right\} + \underbrace{[\mu_T + e \sum_{a \neq 0} \gamma_T^a] \bar{\theta}}_{\Delta_T \hat{e}} \right) \quad (35)
\end{aligned}$$

The function in (35) is linear in x_0, θ_1 and $\bar{\theta}$. So, $L_T x_1$ is linear. By definition of the new coefficient sequence $(\hat{c}, \hat{d}, \hat{e})$ in (35), each element of the new sequence is nonnegative since each element of the original one was so and the new elements are positive weighted sums of the original ones. The sum of the coefficients inside the parentheses on the right hand side of (34) is 1 since g_1 has the convex combination form. Consequently, the total sum of coefficients on the right hand side of (34) is $\Delta_T^{-1}(\alpha_1 + \alpha_2 + \sum_{a \neq 0} \gamma_T^a + \mu_T) = 1$, which proves property (i). The final form in (35) is just a regrouping of elements in (34). The proof of the properties (ii), (iii), and (iv) follows straightforward modifications of the arguments in Lemma 1. Thus, the unique fixed point x_1^* should lie in the set B_G with an associated equilibrium Markovian policy function g_T^{**} .

Therefore, when the symmetric continuation equilibrium policies are Markovian, i.e., $g : \mathbf{X} \times \Theta \times \{1, \dots, T-1\} \mapsto X$, after any history (x^{t-1}, θ^t) , the unique symmetric equilibrium policy in the first period, g_T^{**} is Markovian too. Since the choice of t was arbitrary, this must be true for any length history. Now, construct the policy function g^* as $g_T^*(x_0, \theta_1) = g_T^{**}(x_0, \theta_1)$ for any initial (x_0, θ_1) ; and $g_{T-(t-1)}^*(x_{t-1}, \theta_t) = g_{T-(t-1)}(x_{t-1}, \theta_t)$, for all $t \in \{2, \dots, T\}$ and all (x_{t-1}, θ_t) . But then, the function g^* is by construction the unique MPE of the T -period economy. This completes the induction step for any given $T \geq 2$. Therefore, the claim in Theorem 1 is true for any finite horizon economy.

Step 3: Convergence and stationarity. This step proves that the sequence of finite horizon symmetric Markovian equilibria tends to a stationary symmetric Markov Perfect equilibrium. To do that, we treat finite-horizon economies as finite truncations of the infinite-horizon economy. Let $G^\infty := \prod_{t=1}^\infty G$ be the infinite-horizon Markovian strategy set. For a fixed discount factor $\beta \in (0, 1)$, let $L_\beta := \{\beta_T \in [0, 1]^\infty \mid \beta_{T,t} = \beta^{t-1}, \text{ for } 1 \leq t \leq T, \text{ and } \beta_{T,t} = 0, \text{ for } t > T, \text{ where } T \in \{1, 2, \dots\} \cup \{\infty\}\}$ be the space of exponentially declining sequences (at the rate β) that are equal to zero after the T -th element. Endow L_β with the sup norm.

Lemma 4 (Compactness) L_β and G endowed with the supnorm are compact metric spaces.

Given $g \in G^\infty$, let $x^a(g)$ be agent a 's strategy induced by g , i.e., $x^a(g)(x^{t-1}, \theta^t) = g_t(R^a x_{t-1}, R^a \theta_t)$, for all $a \in \mathbb{A}$ and all (x^{t-1}, θ^t) . Define the objective function U for agent 0 in the class of truncated economies as $U : G^\infty \times L_\beta \times G^\infty$ as

$$U(g^0 ; \beta_T, g) := E \left[\sum_{t=1}^{\infty} \beta_{T,t} u(x_{t-1}^0(g^0), x_t^0(g^0), \{x_t^b(g)\}_{b \in N(0)}, \theta_t^0) \mid (x_0, \theta_1) \right]$$

where u represents the conformity preferences and $N(0) = \{-1, 1\}$ as in Assumption 1. Let the feasibility correspondence $\Gamma : L_\beta \times G^\infty \rightarrow G^\infty$ be defined for $T < \infty$ as $\Gamma(\beta_T, g) = \{g^0 \in G^\infty \mid g_t^0(x, \theta) = \bar{\theta}, \forall t > T, \forall (x, \theta) \in \mathbf{X} \times \Theta\}$, and for $T = \infty$ as $\Gamma(\beta_\infty, g) = G^\infty$. It is easy to see, thanks to the monotonicity of Γ in T (through β_T) and the compactness of G that Γ is a compact-valued and continuous correspondence. Moreover, as the next Lemma shows, the parameterized objective function U is continuous in g^0 , the choice variable.

Lemma 5 (Continuity) For any given $(\beta_T, g) \in L_\beta \times G^\infty$, $U(\cdot ; \beta_T, g)$ is continuous on $\Gamma(\beta_T, g)$ with respect to the product topology.

For every T -period symmetric Markovian equilibrium policy sequence g^{**T} , define $g^{**T} \in G^\infty$ as

$$\forall t, \forall (x, \theta) \in \mathbf{X} \times \Theta, g_t^{**T}(x, \theta) := \begin{cases} g_{T-(t-1)}^{*T}(x, \theta), & \text{if } t \leq T \\ \bar{\theta}, & \text{if } t > T \end{cases}$$

G^∞ endowed with the product topology is compact since each G endowed with the supnorm is compact from Lemma 4. Since product topology is metrizable, say with metric d , (G^∞, d) is a compact metric space hence the sequence $(g^{**T})_T$ has a convergent subsequence $(g^{**T_n})_{T_n}$ in G^∞ that converges say to $g^* \in G^\infty$.⁶⁴ Let $M : L_\beta \times G^\infty \rightarrow G^\infty$ be the correspondence of maximizers of U given the value of the parameters. Also, let $\mathcal{E} : L_\beta \rightarrow G^\infty$ be the symmetric equilibrium correspondence for the sequence of finite economies. Since g^{*T_n} is a symmetric Markovian equilibrium for any T_n , for all $g^{T_n} \in G^\infty$ we have

$$\begin{aligned} U(g_{T_n}^* ; \beta_{T_n}, g_{T_n}^*) &= E \left[\sum_{t=1}^{\infty} \beta_{T_n,t} u(x_{t-1}^0(g^{*T_n}), x_t^0(g^{*T_n}), \{x_t^b(g^{*T_n})\}_{b \in N(0)}, \theta_t^0) \mid (x_0, \theta_1) \right] \\ &= E \left[\sum_{t=1}^{T_n+1} \beta^{t-1} u(x_{t-1}^0(g^{*T_n}), x_t^0(g^{*T_n}), \{x_t^b(g^{*T_n})\}_{b \in N(0)}, \theta_t^0) \mid (x_0, \theta_1) \right] \\ &\geq E \left[\sum_{t=1}^{T_n+1} \beta^{t-1} u(x_{t-1}^0(g^{T_n}), x_t^0(g^{T_n}), \{x_t^b(g^{*T_n})\}_{b \in N(0)}, \theta_t^0) \mid (x_0, \theta_1) \right] \\ &= E \left[\sum_{t=1}^{\infty} \beta_{T_n,t} u(x_{t-1}^0(g^{T_n}), x_t^0(g^{T_n}), \{x_t^b(g^{*T_n})\}_{b \in N(0)}, \theta_t^0) \mid (x_0, \theta_1) \right] \\ &= U(g_{T_n} ; \beta_{T_n}, g_{T_n}^*) \end{aligned}$$

Thus, $g_{T_n}^* \in M(\beta_{T_n}, g_{T_n}^*)$ for all T_n . Since U is continuous in the choice dimension due to Lemma 5 and that the feasibility correspondence Γ is continuous, by the Maximum Theorem (see Berge (1963), p. 115), the

⁶⁴See the proof of Lemma 5 in Appendix D for a metrization product topology.

correspondence of maximizers, M , is upper hemi-continuous. This implies that if $(\beta_{T_n}, g_{T_n}^*) \rightarrow (\beta_\infty, g^*)$, then $g^* \in M(\beta_\infty, g^*)$ hence g^* is a symmetric MPE of the infinite-horizon economy. The immediate implication of this is that the equilibrium correspondence $\mathcal{E}$ is upper hemi-continuous too. Since, each finite-horizon T -period economy has a unique symmetric MPE, $\mathcal{E}$ is single-valued hence continuous for $T < \infty$. Define $\mathcal{F}(\beta_T) := \mathcal{E}(\beta_T)$, for $T < \infty$ and let $\mathcal{F}(\beta_\infty) = g^*$. With this construction, $\mathcal{F}$ is continuous on the space L_β , which is compact under the supnorm by Lemma 4. This makes $\mathcal{F}$ uniformly continuous. So, for a given $\epsilon > 0$, we can pick $\delta > 0$ small enough so that $\|\beta^T - \beta^{T'}\|_\infty < \delta$ implies $d(\mathcal{F}(\beta_T), \mathcal{F}(\beta_{T'})) < \frac{\epsilon}{2}$. We know from the previous approximation that for $\beta_T \rightarrow \beta_\infty$ there is a subsequence $g^{*T_n} \rightarrow g^*$. Since $(\beta_T)_T$ is convergent, it is Cauchy. So, choose $T(\delta)$ large enough such that $\forall T, T' \geq T(\delta)$, $\|\beta_T - \beta_{T'}\| < \delta$ and $\forall T_n \geq T(\delta)$, $\|g^{*T_n} - g^*\|_\infty < \frac{\epsilon}{2}$. Pick an element, T_n , of the subsequence and any other element, T' , such that $T_n, T' \geq T(\delta)$. We have

$$\begin{aligned} d(g^{*T'}, g^*) &= d(\mathcal{F}(\beta_{T'}), \mathcal{F}(\beta_\infty)) \\ &\leq d(\mathcal{F}(\beta_{T'}), \mathcal{F}(\beta_{T_n})) + d(\mathcal{F}(\beta_{T_n}), \mathcal{F}(\beta_\infty)) \\ &< \frac{\epsilon}{2} + d(g^{*T_n}, g^*) \\ &< \epsilon \end{aligned}$$

The first inequality is the triangle inequality; the second is due to the uniform continuity of $\mathcal{F}$ and the third is by the fact that $g^{*T_n} \rightarrow g^*$ uniformly. This proves that the whole sequence $g^{*T} \rightarrow g^*$ uniformly. The implication of this latter is that, as the finite-horizon economies approach the infinite-horizon economy, every two consecutive period, we make choices approximately with respect to the same MPE policy, hence g^* is stationary. This concludes **Step 3** which in turn establishes the proof of the statement of Theorem 1. ■

11 Appendix B: Proof of Inefficiency (Theorem 5)

We give the proof for economies with complete information. Once again, the extension of the line of proof to the incomplete information economies is straightforward.

Finite-Horizon

Take any finite horizon economy ($T < \infty$). We will use continuity arguments so endow the underlying space $\mathbf{X} \times \Theta$ with the product topology. Product topology is metrizable, say by metric d^{65} . In the final period of this finite horizon economy, with absolutely continuous distribution π_{T-1} on the space of choice profiles x_{T-1} ⁶⁶ with a positive density, the planner maximizes ex-ante (before the realization of θ_T) the expected

⁶⁵Let $|\cdot|$ be the usual Euclidean norm. For any $(x, \theta), (x', \theta') \in \mathbf{X} \times \Theta$, let

$$d((x, \theta), (x', \theta')) := \sum_{a \in A} 2^{-a} (|x_a - x'_a| + |\theta_a - \theta'_a|)$$

Since $X = \Theta = [\underline{x}, \bar{x}]$ is a compact interval, this is a well-defined metric that metrizes the product topology on $\mathbf{X} \times \Theta$. See also Aliprantis and Border (2006, p. 90)

⁶⁶Stating with an initial π_0 which is absolutely continuous, the MPE policy function and the absolutely continuous

utility of a given agent, say of agent $0 \in \mathbb{A}$, by choosing a symmetric policy function $h \in CB(\mathbf{X} \times \Theta, X)$, the space of bounded, continuous, and X -valued measurable functions.⁶⁷

$$\max_{\{h \in CB(\mathbf{X} \times \Theta, X)\}} \int u(x_{T-1}^0, h(x_{T-1}, \theta_T), h(R^{-1}x_{T-1}, R^{-1}\theta_T), h(Rx_{T-1}, R\theta_T), \theta_T^0) \mathbb{P}(d\theta_T) \pi_{T-1}(dx_{T-1})$$

The space $\mathbf{X} \times \Theta$ is compact with respect to the product topology since X and Θ are compact. Since the utility function is continuous and strictly concave in all arguments, the maximizer exists and it is unique. The necessary condition for optimality is summarized in the following lemma.

Lemma 6 *The first order necessary condition for optimality requires that, for any $(x_{T-1}, \theta_T) \in \mathbf{X} \times \Theta$ ⁶⁸*

$$\begin{aligned} 0 &= \alpha_1 (x_{T-1}^0 - h(x_{T-1}, \theta_T)) + \alpha_2 (\theta_T^0 - h(x_{T-1}, \theta_T)) \\ &+ \alpha_3 (h(R^{-1}x_{T-1}, R^{-1}\theta_T) - h(x_{T-1}, \theta_T)) + \alpha_3 (h(Rx_{T-1}, R\theta_T) - h(x_{T-1}, \theta_T)) \\ &- \alpha_3 (h(x_{T-1}, \theta_T) - h(Rx_{T-1}, R\theta_T)) \\ &- \alpha_3 (h(x_{T-1}, \theta_T) - h(R^{-1}x_{T-1}, R^{-1}\theta_T)) \end{aligned}$$

This implies that

$$h(x_{T-1}, \theta_T) [\alpha_1 + \alpha_2 + 4\alpha_3] = \alpha_1 x_{T-1}^0 + \alpha_2 \theta_T^0 + 2\alpha_3 h(R^{-1}x_{T-1}, R^{-1}\theta_T) + 2\alpha_3 h(Rx_{T-1}, R\theta_T) \quad (36)$$

Following the proof of existence, note that the operator induced by equation (36) is a contraction on the Banach space of bounded, continuous, measurable functions with the supnorm, whose unique fixed point turns out to be in G , the space of linear policy maps that have the convex combination form, defined in (25). Therefore, one can fit the following solution

$$h(x_{T-1}, \theta_T) = \sum_a c_P^a x_{T-1}^a + \sum_a d_P^a \theta_T^a$$

substituting, we get

$$\begin{aligned} \sum_a c_P^a x_{T-1}^a + \sum_a d_P^a \theta_T^a &= (\alpha_1 + \alpha_2 + 4\alpha_3)^{-1} \left[\alpha_1 x_{T-1}^0 + \alpha_2 \theta_T^0 \right. \\ &+ 2\alpha_3 \left(\sum_a c_P^a x_{T-1}^{a-1} + \sum_a d_P^a \theta_T^{a-1} \right) \\ &\left. + 2\alpha_3 \left(\sum_a c_P^a x_{T-1}^{a+1} + \sum_a d_P^a \theta_T^{a+1} \right) \right] \end{aligned}$$

By matching coefficients, we get

$$c_P^a = (\alpha_1 + \alpha_2 + 4\alpha_3)^{-1} \left[2\alpha_3 c_P^{a-1} + 2\alpha_3 c_P^{a+1} + \alpha_1 \mathbf{1}_{\{a=0\}} \right], \quad \forall a \in \mathbb{A}$$

preference shocks will induce a sequence (π_t) of absolutely continuous distributions on t -period equilibrium choice profiles.

⁶⁷Since the planner's choice rule is symmetric, the choice of agent 0 rather than another agent is inconsequential.

⁶⁸The proof is in Appendix D.

and

$$d_P^a = (\alpha_1 + \alpha_2 + 4\alpha_3)^{-1} \left[2\alpha_3 d_P^{a-1} + 2\alpha_3 d_P^{a+1} + \alpha_2 \mathbf{1}_{\{a=0\}} \right], \quad \forall a \in \mathbb{A}$$

One can solve for the coefficient sequence in closed form by mimicking the same proof that we provided to characterize the equilibrium policy function weights in Theorem 3. Thus, one gets for any $a \in \mathbb{A}$,

$$c_P^a = r_P^{|a|} \left(\frac{\alpha_1}{\alpha_1 + \alpha_2} \right) \left(\frac{1 - r_P}{1 + r_P} \right) \quad \text{and} \quad d_P^a = r_P^{|a|} \left(\frac{\alpha_2}{\alpha_1 + \alpha_2} \right) \left(\frac{1 - r_P}{1 + r_P} \right) \quad (37)$$

where

$$r_P = \left(\frac{\Delta_P}{2\alpha_3} \right) - \sqrt{\left(\frac{\Delta_P}{2\alpha_3} \right)^2 - 1} \quad \text{and} \quad \Delta_P = \alpha_1 + \alpha_2 + 4\alpha_3$$

We next compare the equilibrium policy sequence in Theorem 3 with the planner's optimal choice coefficient sequence. Notice that

$$\left(\frac{\Delta_P}{2\alpha_3} \right) = \frac{\alpha_1 + \alpha_2 + 4\alpha_3}{2\alpha_3} = 2 + \frac{\alpha_1 + \alpha_2}{2\alpha_3} < 2 + \frac{\alpha_1 + \alpha_2}{\alpha_3} = \left(\frac{\Delta_1}{\alpha_3} \right)$$

which implies from (51) that $r_P > r_1$. Thus, the planner's optimal policy coefficient sequence converges to zero slower than the equilibrium policy coefficient sequence. This also means that the equilibrium policy cannot satisfy the FOC of the planner's problem. Therefore, the equilibrium is inefficient for finite-horizon economies.

Infinite-Horizon

The argument here is very similar to the one in the finite horizon case. We know from Theorem 1 that the equilibrium whose existence we are assured has the following structure

$$g(x_{T-1}, \theta_T) = \sum_a c^a x_{T-1}^a + \sum_a d^a \theta_T^a + e \bar{\theta}$$

We argue that this solution cannot satisfy the planner's problem's optimality condition. For a given function $h \in G$ (see (25)), define $H : \mathbf{X} \times \Theta \rightarrow \mathbf{X}$ as

$$H(x_{T-1}, \theta_T) := (h(R^a x_{T-1}, R^a \theta_T))_{a \in \mathbb{A}} \quad (38)$$

Let V^h be the continuation value of using the function h in the future, defined recursively as

$$\begin{aligned} V^h(x_{T-1}, \theta_T) &= u(x_{T-1}^0, h(x_{T-1}, \theta_T), h(R^{-1} x_{T-1}, R^{-1} \theta_T), h(R x_{T-1}, R \theta_T), \theta_T^0) \\ &\quad + \beta \int V^h(H(x_{T-1}, \theta_T), \theta_{T+1}) \mathbb{P}(d\theta_{T+1}) \end{aligned} \quad (39)$$

where u represents the conformity preferences in Assumption 1. Since the policy $h \in G$ is linear and the utility function is continuously differentiable and strictly concave with respect to all arguments, elementary dynamic programming techniques (see for e.g. Stokey and Lucas (1989)) guarantee that for the given choice rule $h \in G$, the value function V^h exists, it is bounded, continuous, strictly concave and continuously differentiable. We will also denote by V_a^h the partial derivative of V^h with respect to agent a 's initial choice. Given an initial absolutely continuous distribution π_{T-1} on the space of previous period's choice

profiles with a positive density, the planner maximizes ex-ante (before the realization of θ_T) the expected discounted utility of a given agent, say of agent 0 $\in \mathbb{A}$. So, the planner's problem is

$$\begin{aligned} \max_{\{h \in G\}} \quad & \int \left[u(x_{T-1}^0, h(x_{T-1}, \theta_T), h(R^{-1} x_{T-1}, R^{-1} \theta_T), h(R x_{T-1}, R \theta_T), \theta_T^0) \right. \\ & \left. + \beta \int V^h(H(x_{T-1}, \theta_T), \theta_{T+1}) \right] \mathbb{P}(d\theta_T) \mathbb{P}(d\theta_{T+1}) \pi_{T-1}(dx_{T-1}) \end{aligned}$$

As in the finite case, the solution to this problem exists and it is unique thanks to the compactness (with respect to the product topology) of the underlying space $\mathbf{X} \times \Theta$ and the continuity and strict concavity of the utility and value functions. A straightforward modification of the first order condition argument in the finite case yields, for any $(x_{T-1}, \theta_T) \in \mathbf{X} \times \Theta$

$$\begin{aligned} 0 &= u_2(x_{T-1}^0, h(x_{T-1}, \theta_T), h(R^{-1} x_{T-1}, R^{-1} \theta_T), h(R x_{T-1}, R \theta_T), \theta_T^0) \\ &+ u_3(x_{T-1}^1, h(R x_{T-1}, R \theta_T), h(x_{T-1}, \theta_T), h(R^2 x_{T-1}, R^2 \theta_T), \theta_T^1) \\ &+ u_4(x_{T-1}^{-1}, h(R^{-1} x_{T-1}, R^{-1} \theta_T), h(R^{-2} x_{T-1}, R^{-2} \theta_T), h(x_{T-1}, \theta_T), \theta_T^{-1}) \\ &+ \beta \sum_a \int V_a^h(H(R^{-a} x_{T-1}, R^{-a} \theta_T), \theta_{T+1}) \mathbb{P}(d\theta_{T+1}) \end{aligned}$$

But if the equilibrium policy is h , the FOC yields

$$\begin{aligned} 0 &= \int u_2(x_{T-1}^0, h(x_{T-1}, \theta_T), h(R^{-1} x_{T-1}, R^{-1} \theta_T), h(R x_{T-1}, R \theta_T), \theta_T^0) \\ &+ \beta \int V_a^h(H(x_{T-1}, \theta_T), \theta_{T+1}) \mathbb{P}(d\theta_{T+1}) \end{aligned}$$

For the same solution to satisfy both FOCs, it has to be the case that for any $(x_{T-1}, \theta_T) \in \mathbf{X} \times \Theta$, the difference of the two FOCs is zero, i.e.

$$\begin{aligned} 0 &= u_3(x_{T-1}^1, h(R x_{T-1}, R \theta_T), h(x_{T-1}, \theta_T), h(R^2 x_{T-1}, R^2 \theta_T), \theta_T^1) \\ &+ u_4(x_{T-1}^{-1}, h(R^{-1} x_{T-1}, R^{-1} \theta_T), h(R^{-2} x_{T-1}, R^{-2} \theta_T), h(x_{T-1}, \theta_T), \theta_T^{-1}) \\ &+ \beta \sum_{a \neq 0} \int V_a^h(H(R^{-a} x_{T-1}, R^{-a} \theta_T), \theta_{T+1}) \mathbb{P}(d\theta_{T+1}) \end{aligned}$$

For the quadratic specification, this entails

$$\begin{aligned} 0 &= 2\alpha_3 (h(x_{T-1}, \theta_T) - h(R x_{T-1}, R \theta_T)) + 2\alpha_3 (h(x_{T-1}, \theta_T) - h(R^{-1} x_{T-1}, R^{-1} \theta_T)) \\ &+ \beta \sum_{a \neq 0} \int V_a^h(H(R^{-a} x_{T-1}, R^{-a} \theta_T), \theta_{T+1}) \mathbb{P}(d\theta_{T+1}) \end{aligned}$$

Substituting the equilibrium policy function g and recollecting terms

$$\begin{aligned} & 2\alpha_3 \left[\sum_a c^a (2x_{T-1}^a - x_{T-1}^{a-1} - x_{T-1}^{a+1}) + \sum_a d^a (2\theta_{T-1}^a - \theta_{T-1}^{a-1} - \theta_{T-1}^{a+1}) \right] \\ & + \beta \sum_{a \neq 0} \int V_a^h(H(R^{-a} x_{T-1}, R^{-a} \theta_T), \theta_{T+1}) \mathbb{P}(d\theta_{T+1}) = 0 \end{aligned} \quad (40)$$

We next show in the following lemma that there exists a positive measure subset of the underlying space on which the above expression assumes non-zero values.

Lemma 7 Let $(\hat{x}, \hat{\theta}) \in \mathbf{X} \times \Theta$ be the point where $\hat{x}_a = \bar{x}$ and $\hat{\theta}_a = \bar{x}$, for all $a \in \mathbb{A}$ ⁶⁹. Then the expression in (40) is negative on a positive measure subset $E \subset \mathbf{X} \times \Theta$, that includes $(\hat{x}, \hat{\theta})$.

The statement of Lemma 7 leads to a contradiction since it means that the planner's optimal rule and the equilibrium policy function g does not agree on E . Therefore, g is inefficient. This concludes the proof. ■

12 Appendix C: Proof of Ergodicity (Theorem 4)

Suppose that the process $((\theta_t^a)_{t=-\infty}^\infty)_{a \in \mathbb{A}}$ is i.i.d. with respect to a and t according to ν . Let π be the initial measure on the configuration space $\mathbf{X}$ which is the distribution of

$$x_0 = \left(\frac{e\bar{\theta}}{1-C} + \sum_{s=1}^\infty \sum_{b_1} \dots \sum_{b_s} c^{b_1} \dots c^{b_{s-1}} \left(d^{b_s} \theta_{1-s}^{a+b_1+\dots+b_s} \right) \right)_{a \in \mathbb{A}} \quad (41)$$

Given that $(x_t \in \mathbf{X})_{t=0}^\infty$ is an equilibrium process generated by the stationary MPE g^* in Theorem 1, given x_0 , one obtains on the equilibrium path

$$\begin{aligned} x_1^a &= \sum_{b_1 \in \mathbb{A}} c^{b_1} x_0^{a+b_1} + \sum_{b_1 \in \mathbb{A}} d^{b_1} \theta_1^{a+b_1} + e\bar{\theta} \\ &= \sum_{b_1 \in \mathbb{A}} c^{b_1} \left(\frac{e\bar{\theta}}{1-C} + \sum_{s=1}^\infty \sum_{b_1} \dots \sum_{b_s} c^{b_1} \dots c^{b_{s-1}} \left(d^{b_s} \theta_{1-s}^{a+b_1+\dots+b_s} \right) \right) \\ &\quad + \sum_{b_1 \in \mathbb{A}} d^{b_1} \theta_1^{a+b_1} + e\bar{\theta} \\ &= \sum_{s=1}^\infty \sum_{b_1} \dots \sum_{b_{s+1}} c^{b_1} \dots c^{b_s} \left(d^{b_{s+1}} \theta_{1-s}^{a+b_1+\dots+b_{s+1}} \right) + \sum_{b_1 \in \mathbb{A}} d^{b_1} \theta_1^{a+b_1} \\ &\quad + C \frac{e\bar{\theta}}{1-C} + e\bar{\theta} \\ &= \frac{e\bar{\theta}}{1-C} + \sum_{s=1}^\infty \sum_{b_1} \dots \sum_{b_s} c^{b_1} \dots c^{b_{s-1}} \left(d^{b_s} \theta_{2-s}^{a+b_1+\dots+b_s} \right) \end{aligned}$$

which has exactly the same form as in (41). Hence, x_0^a and x_1^a are distributed identically when the initial measure is π . Since the choice of a was arbitrary, π is a stationary distribution of the Markov process $(x_t)_{t=0}^\infty$. Moreover, from Lemma 2 for a stationary policy function, on any arbitrary path $(\theta_1, \theta_2, \dots)$ of

⁶⁹Recall from Assumption 1 that $\bar{x}$ is the upper boundary of the feasible action and type sets X and Θ

the stochastic process

$$\begin{aligned}
x_t^a &= \sum_{b_1} \cdots \sum_{b_t} c^{b_1} \cdots c^{b_t} x_0^{a+b_1+\cdots+b_t} \\
&\quad + \sum_{s=1}^t \sum_{b_1} \cdots \sum_{b_s} c^{b_1} \cdots c^{b_{s-1}} \left(d^{b_s} \theta_{t-(s-1)}^{a+b_1+\cdots+b_s} + e \bar{\theta} \right) \\
&= C^t \sum_{b_1} \cdots \sum_{b_t} \left(\frac{c^{b_1} \cdots c^{b_t}}{C^t} \right) x_0^{a+b_1+\cdots+b_t} \\
&\quad + \sum_{s=1}^t \sum_{b_1} \cdots \sum_{b_s} c^{b_1} \cdots c^{b_{s-1}} \left(d^{b_s} \theta_{t+1-s}^{a+b_1+\cdots+b_s} + e \bar{\theta} \right)
\end{aligned} \tag{42}$$

Thus, independent of the initial conditions, x_t^a converges pointwise to $x^a \in X$, i.e.,

$$\begin{aligned}
x^a := \lim_{t \rightarrow \infty} x_t^a &= \lim_{t \rightarrow \infty} \left[C^t \sum_{b_1} \cdots \sum_{b_t} \left(\frac{c^{b_1} \cdots c^{b_t}}{C^t} \right) x_0^{a+b_1+\cdots+b_t} \right. \\
&\quad \left. + \sum_{s=1}^t \sum_{b_1} \cdots \sum_{b_s} c^{b_1} \cdots c^{b_{s-1}} \left(d^{b_s} \theta_{t+1-s}^{a+b_1+\cdots+b_s} + e \bar{\theta} \right) \right]
\end{aligned} \tag{43}$$

The first term of the previous expression $C^t \rightarrow 0$ since $C < 1$ due to $\alpha_i > 0$, for all i . The first term in the parentheses in the summand is a convex combination of uniformly bounded terms. Hence, the first part of the above expression goes to 0 as $t \rightarrow \infty$. Moreover, since the equilibrium is symmetric, the convergence is uniform across agents: $x_t \rightarrow x = (x^a)$ uniformly. Since the exogenous shock process is i.i.d, the part after the plus sign in (43) is identical to the distribution of

$$\sum_{s=1}^t \sum_{b_1} \cdots \sum_{b_s} c^{b_1} \cdots c^{b_{s-1}} \left(d^{b_s} \theta_{1-s}^{a+b_1+\cdots+b_s} + e \bar{\theta} \right)$$

which is the ‘ t -translated-into-the-past’ version of the former. Thus, for any given initial value x_0 , and a path $(\dots, \theta_{-1}, \theta_0)$, the pointwise limit of x_t^a can be written as

$$x^a = \frac{e \bar{\theta}}{1 - C} + \sum_{s=1}^{\infty} \sum_{b_1} \cdots \sum_{b_s} c^{b_1} \cdots c^{b_{s-1}} \left(d^{b_s} \theta_{1-s}^{a+b_1+\cdots+b_s} \right) \tag{44}$$

For the rest of the proof, let $\mathbb{P}_{\infty}(\cdot) := \prod_{t=0}^{\infty} \mathbb{P}(\cdot)$ and $\theta := (\dots, \theta_{-1}, \theta_0, \cdot)$. We next show that, for any arbitrary initial distribution π_0 , the sequence of equilibrium distributions π_t generated by the exogenous law $\mathbb{P}$ and the stationary MPE policy g^* converges weakly to the invariant distribution π . To that effect, pick any $f \in C(\mathbf{X}, \mathbb{R})$, the set of bounded, continuous, and measurable, real-valued functions from $\mathbf{X}$ into

$\mathbb{R}$. Let π_0 be an arbitrary initial distribution for x_0 . We have

$$\begin{aligned}
\lim_{t \rightarrow \infty} \int f(x_t) \pi_t(dx_t) &= \lim_{t \rightarrow \infty} \int f \left(\left(C^t \sum_{b_1} \cdots \sum_{b_t} \left(\frac{c^{b_1} \cdots c^{b_t}}{C^t} \right) x_0^{a+b_1+\cdots+b_t} \right. \right. \\
&\quad \left. \left. \sum_{s=1}^t \sum_{b_1} \cdots \sum_{b_s} c^{b_1} \cdots c^{b_{s-1}} \left(d^{b_s} \theta_{1-s}^{a+b_1+\cdots+b_s} + e \bar{\theta} \right) \right) \right)_{a \in \mathbb{A}} \right) \mathbb{P}_\infty(d\theta) \pi_0(dx_0) \\
&= \int f \left(\left(\frac{e \bar{\theta}}{1-C} + \sum_{s=1}^\infty \sum_{b_1} \cdots \sum_{b_s} c^{b_1} \cdots c^{b_{s-1}} \left(d^{b_s} \theta_{1-s}^{a+b_1+\cdots+b_s} \right) \right) \right)_{a \in \mathbb{A}} \right) \mathbb{P}_\infty(d\theta) \pi_0(dx_0) \\
&= \int f(x) \pi(dx)
\end{aligned} \tag{45}$$

The first equality is from (42); the second is due to Lebesgue Dominated Converge theorem; third is due to the continuity of f and the pointwise limit of x_t in (44). Thus, for any $f \in C(\mathbf{X}, \mathbb{R})$, $\lim_{t \rightarrow \infty} \int f d\pi_t = \int f d\pi$, meaning that the sequence π_t converges weakly to π , the invariant distribution of the equilibrium process. The choice of π_0 was arbitrary. Hence, for any initial distribution, the process induced converges weakly to the same invariant distribution π . Therefore, π is the unique invariant distribution of the equilibrium process. Here is why: Suppose that $\hat{\pi}$ is another invariant distribution. This implies that the induced process starting with $\pi_0 = \hat{\pi}$ should satisfy $\pi_t = \hat{\pi}$, for all $t = 1, 2, \dots$. From the above convergence argument $\pi_t \rightarrow \pi$ weakly. Hence $\hat{\pi} = \pi$.

Finally, to show ergodicity, pick an $f \in B(\mathbf{X}, \mathbb{R})$, the set of bounded, measurable, real-valued functions from $\mathbf{X}$ into $\mathbb{R}$. The process starting with π is stationary, hence $\pi_t = \pi$ for all $t = 0, 1, \dots$. Since the process x_t is stationary, so is the process $(f(x_t))$. We can then use Birkhoff's Ergodic Theorem (see e.g. Aliprantis and Border (2006), p. 659) on the process $(f(x_t))$ to obtain

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T f(x_t) = \int f(x_t) \pi(dx_t)$$

almost surely. Since the choice of f was arbitrary, the last expression holds for all $f \in B(\mathbf{X}, \mathbb{R})$. Thus the equilibrium process $(x_t \in \mathbf{X})_{t=0}^\infty$ starting from initial distribution π is ergodic. This concludes the proof of Theorem 4. ■

13 Appendix D: Technical Appendix

Proof of Lemma 2 :(i) This part is simply by iterated application of the policy maps, i.e.,

$$\begin{aligned}
x_t^a &= g_{T-(t-1)}(R^a x_{t-1}, R^a \theta_t) \\
&= \sum_{b_1 \in \mathbb{A}} c_{T-(t-1)}^{b_1} x_{t-1}^{a+b_1} + \sum_{b_1 \in \mathbb{A}} d_{T-(t-1)}^{b_1} \theta_t^{a+b_1} + e_{T-(t-1)} \bar{\theta} \\
&= \sum_{b_1 \in \mathbb{A}} c_{T-(t-1)}^{b_1} \underbrace{g_{T-t}(R^{a+b_1} x_{t-2}, R^{a+b_1} \theta_{t-1})}_{x_{t-1}^{a+b_1}} + \sum_{b_1 \in \mathbb{A}} d_{T-(t-1)}^{b_1} \theta_t^{a+b_1} + e_{T-(t-1)} \bar{\theta} \\
&= \sum_{b_1 \in \mathbb{A}} c_{T-(t-1)}^{b_1} \left(\sum_{b_2 \in \mathbb{A}} c_{T-(t-2)}^{b_2} x_{t-2}^{a+b_1+b_2} + \sum_{b_2 \in \mathbb{A}} d_{T-(t-2)}^{b_2} \theta_{t-1}^{a+b_1+b_2} + e_{T-(t-2)} \bar{\theta} \right) \\
&\quad + \sum_{b_1 \in \mathbb{A}} d_{T-(t-1)}^{b_1} \theta_t^{a+b_1} + e_{T-(t-1)} \bar{\theta} \\
&\vdots \\
&= \sum_{b_1 \in \mathbb{A}} \cdots \sum_{b_{t-1} \in \mathbb{A}} c_{T-(t-1)}^{b_1} \cdots c_{T-1}^{b_{t-1}} x_1^{a+b_1+\cdots+b_{t-1}} \\
&\quad + \sum_{s=1}^{t-1} \sum_{b_1 \in \mathbb{A}} \cdots \sum_{b_{s-1} \in \mathbb{A}} c_{T-(t-1)}^{b_1} \cdots c_{T-(s)+1}^{b_{s-1}} \left(\sum_{b_s \in \mathbb{A}} d_{T-(t-s)}^{b_s} \theta_{t-(s-1)}^{a+b_1+\cdots+b_s} + e_{T-(t-s)} \bar{\theta} \right)
\end{aligned}$$

(ii) For $t = 2$, $\frac{\partial}{\partial x_1^a} x_2^0 = c_1^a \geq c_1^b = \frac{\partial}{\partial x_1^b} x_2^0$ by (25) (iii). Suppose the claim is true for $t \leq k$ and let $t = k + 1$. Assume w.l.o.g that $a < b$. Let $\underline{s} := \max\{s \in \mathbb{A} : s \leq \frac{a+b}{2}\}$ and $\bar{s} := \min\{s \in \mathbb{A} : s \geq \frac{a+b}{2}\}$. This implies that $\frac{\partial}{\partial x_1^a} x_k^s - \frac{\partial}{\partial x_1^b} x_k^s \geq 0$ (≤ 0) for $s \leq \underline{s}$ ($s \geq \bar{s}$). Due to the assumed symmetry, $\left[\frac{\partial}{\partial x_1^a} x_k^{s-s} - \frac{\partial}{\partial x_1^b} x_k^{s-s} \right] = \left[\frac{\partial}{\partial x_1^{a-\underline{s}+s}} x_k^0 - \frac{\partial}{\partial x_1^{b-\underline{s}+s}} x_k^0 \right]$ and $\left[\frac{\partial}{\partial x_1^{a-\bar{s}-s}} x_k^0 - \frac{\partial}{\partial x_1^{b-\bar{s}-s}} x_k^0 \right] = \left[\frac{\partial}{\partial x_1^a} x_k^{\bar{s}} - \frac{\partial}{\partial x_1^b} x_k^{\bar{s}} \right]$. This implies that for any $s > 0$

$$\left[\frac{\partial}{\partial x_1^{a-\underline{s}+\tau}} x_k^0 - \frac{\partial}{\partial x_1^{b-\underline{s}+\tau}} x_k^0 \right] = - \left[\frac{\partial}{\partial x_1^{a-\bar{s}-\tau}} x_k^0 - \frac{\partial}{\partial x_1^{b-\bar{s}-\tau}} x_k^0 \right]$$

Thus, we can use this to separate $\mathbb{A}$ into $\{s \in \mathbb{A} : s \leq \underline{s}\}$ $\{s \in \mathbb{A} : s \geq \bar{s}\}$ and rearrange the sum

$$\begin{aligned}
\frac{\partial}{\partial x_1^a} x_t^0 - \frac{\partial}{\partial x_1^b} x_t^0 &= \sum_{s \in \mathbb{A}} c_{T-k}^s \left[\frac{\partial}{\partial x_1^a} x_k^s - \frac{\partial}{\partial x_1^b} x_k^s \right] \\
&= \sum_{s \in \mathbb{A}} c_{T-k}^s \left[\frac{\partial}{\partial x_1^{a-s}} x_k^0 - \frac{\partial}{\partial x_1^{b-s}} x_k^0 \right] \\
&= \sum_{\tau \geq 0} \left(c_{T-k}^{\underline{s}-s} - c_{T-k}^{\bar{s}+s} \right) \left[\frac{\partial x_k^0}{\partial x_1^{a-\underline{s}+s}} - \frac{\partial x_k^0}{\partial x_1^{b-\bar{s}+s}} \right] \\
&\geq 0
\end{aligned}$$

The term in the brackets is nonnegative by hypothesis. Since $a < b$, $\underline{s} \geq 0$ which implies that $c_{T-k}^{\underline{s}} \geq c_{T-k}^{\bar{s}}$. But this implies that $c_{T-k}^{\underline{s}-s} \geq c_{T-k}^{\bar{s}+s}$ for any $s \geq 0$ which means that the argument in the parenthesis is nonnegative too. So, the claim is true. The analysis for the case $a > b$ is a straightforward modification of the same argument.

(iii) Using the t -th period equilibrium policy

$$\begin{aligned}
\frac{\partial}{\partial x_1^a} (2x_t^0 - x_t^1 - x_t^{-1}) &= \frac{\partial}{\partial x_1^a} \left[2 \sum_{b \in \mathbb{A}} c_{T-(t-1)}^b x_{t-1}^b + \sum_{b \in \mathbb{A}} d_{T-(t-1)}^b \theta_t^b + e_{T-(t-1)} \bar{\theta} \right. \\
&\quad - \sum_{b \in \mathbb{A}} c_{T-(t-1)}^{b-1} x_{t-1}^b + \sum_{b \in \mathbb{A}} d_{T-(t-1)}^{b-1} \theta_t^b + e_{T-(t-1)} \bar{\theta} \\
&\quad \left. - \sum_{b \in \mathbb{A}} c_{T-(t-1)}^{b+1} x_{t-1}^b + \sum_{b \in \mathbb{A}} d_{T-(t-1)}^{b+1} \theta_t^b + e_{T-(t-1)} \bar{\theta} \right] \\
&= \sum_{b \in \mathbb{A}} \left(2c_{T-(t-1)}^b - c_{T-(t-1)}^{b-1} - c_{T-(t-1)}^{b+1} \right) \frac{\partial}{\partial x_1^a} x_{t-1}^b \leq 0
\end{aligned} \tag{46}$$

The weights in the last parenthesis are negative by property (ii) in (25). By iteratively applying the policy functions from period t backwards, at each iteration the weights on one-period before choices would all be positive and one preserves the convex combination form. This process ends after $t-1$ iteration, the end result being a convex combination of $(x_1^b)_{b \in \mathbb{A}}$, θ^t and $\bar{\theta}$. Thus, the weight on x_1^a is positive, which makes the last term in the last line positive. Therefore the claim is true.

(iv) Let $t \geq 2$.

$$\begin{aligned}
\frac{\partial}{\partial x_1^0} x_t^0 &= \sum_{a \in \mathbb{A}} c_{T-(t-1)}^a \frac{\partial x_{t-1}^a}{\partial x_1^0} = \sum_{a \in \mathbb{A}} c_{T-(t-1)}^a \frac{\partial x_{t-1}^0}{\partial x_1^a} \\
&\leq \sum_{a \in \mathbb{A}} c_{T-(t-1)}^a \frac{\partial x_{t-1}^0}{\partial x_1^0} = C_{T-(t-1)} \frac{\partial x_{t-1}^0}{\partial x_1^0}
\end{aligned}$$

First and second equalities and the first inequality are by the definition of the policy mapping and (i) of Lemma 2; $C_{T-(t-1)}$ is the sum of coefficients on the past history in the period t policy. Since $g_{T-(t-1)}$ satisfies (32), coefficients should match and we should have

$$\begin{aligned}
0 &= c_{T-(t-1)}^a \Delta_{T-(t-1)} - \alpha_1 I_{\{a=0\}} - \sum_{b \neq 0} \gamma_{T-(t-1)}^b c_{T-(t-1)}^{a-b} \\
\text{summing over } a, \quad 0 &= C_{T-(t-1)} \Delta_{T-(t-1)} - \alpha_1 - \sum_{b \neq 0} \gamma_{T-(t-1)}^b C_{T-(t-1)}
\end{aligned}$$

But $\Delta_{T-(t-1)} = \alpha_1 + \alpha_2 + \sum_{b \neq 0} \gamma_{T-(t-1)}^b + \mu_{T-(t-1)}$ by definition. So,

$$C_{T-(t-1)} = \frac{\alpha_1}{\Delta_{T-(t-1)} - \sum_{b \neq 0} \gamma_{T-(t-1)}^b} = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \mu_{T-(t-1)}} \leq \frac{\alpha_1}{\alpha_1 + \alpha_2} \tag{47}$$

Thus,

$$\begin{aligned}
\frac{\partial}{\partial x_1^0} x_t^0 &\leq C_{T-(t-1)} \frac{\partial}{\partial x_1^0} x_{t-1}^0 \\
&\leq \left(\frac{\alpha_1}{\alpha_1 + \alpha_2} \right) \frac{\partial}{\partial x_1^0} x_{t-1}^0
\end{aligned}$$

which proves the claim. ■

Proof of Lemma 3: The coefficient of x_1^a in (32), γ_T^a , is the total effect of a change in x_1^a ($a \neq 0$) on the expected discounted marginal utility of agent 0 (the right hand side of (30)), i.e.,

$$\begin{aligned} \gamma_T^a &:= \alpha_3 I_{\{a \in \{-1, 1\}\}} \\ &\quad - \sum_{\tau=2}^T \beta^{\tau-1} \left(\alpha_1 \frac{\partial}{\partial x_1^a} (x_{\tau-1}^0 - x_\tau^0) \frac{\partial}{\partial x_1^0} (x_{\tau-1}^0 - x_\tau^0) + \alpha_2 \frac{\partial}{\partial x_1^a} x_\tau^0 \frac{\partial}{\partial x_1^0} x_\tau^0 \right. \\ &\quad \left. + \alpha_3 \frac{\partial}{\partial x_1^a} (x_\tau^{-1} - x_\tau^0) \frac{\partial}{\partial x_1^0} (x_\tau^{-1} - x_\tau^0) + \alpha_3 \frac{\partial}{\partial x_1^a} (x_\tau^1 - x_\tau^0) \frac{\partial}{\partial x_1^0} (x_\tau^1 - x_\tau^0) \right) \end{aligned} \quad (48)$$

For any $\tau \geq 2$, the last two term in the summand for each period in equation (48) can be written as

$$\begin{aligned} &\frac{\partial}{\partial x_1^0} (x_\tau^1 - x_\tau^0) \left[\alpha_3 \frac{\partial}{\partial x_1^a} (x_\tau^{-1} - x_\tau^0) + \alpha_3 \frac{\partial}{\partial x_1^a} (x_\tau^1 - x_\tau^0) \right] \\ &= \frac{\partial}{\partial x_1^0} (x_\tau^1 - x_\tau^0) \left[\alpha_3 \frac{\partial x_\tau^0}{\partial x_1^{a+1}} + \alpha_3 \frac{\partial x_\tau^0}{\partial x_1^{a-1}} - 2\alpha_3 \frac{\partial x_\tau^0}{\partial x_1^a} \right] \\ &\leq 0 \end{aligned} \quad (49)$$

The equality is due to the symmetry of the policy function across agents; Lemma 2 (ii) and (iii) imply that the terms in the parentheses are non-positive and the terms in the brackets are non-negative, respectively. Similarly, the first terms in the summand in (48) can be written as

$$\begin{aligned} &\alpha_1 \frac{\partial}{\partial x_1^a} (x_\tau^0 - x_{\tau-1}^0) \frac{\partial}{\partial x_1^0} (x_\tau^0 - x_{\tau-1}^0) + \alpha_2 \frac{\partial}{\partial x_1^a} x_\tau^0 \frac{\partial}{\partial x_1^0} x_\tau^0 \\ &\leq \alpha_1 \frac{\partial}{\partial x_1^a} x_\tau^0 \frac{\partial}{\partial x_1^0} (x_\tau^0 - x_{\tau-1}^0) + \alpha_2 \frac{\partial}{\partial x_1^a} x_\tau^0 \frac{\partial}{\partial x_1^0} x_\tau^0 \\ &= \frac{\partial x_\tau^0}{\partial x_1^a} \left[(\alpha_1 + \alpha_2) \frac{\partial}{\partial x_1^0} x_\tau^0 - \alpha_1 \frac{\partial}{\partial x_1^0} x_{\tau-1}^0 \right] \\ &\leq 0 \end{aligned}$$

which is nonpositive since for any $\tau \geq 2$

$$\frac{\partial}{\partial x_1^0} x_\tau^0 \leq \frac{\alpha_1}{(\alpha_1 + \alpha_2)} \frac{\partial}{\partial x_1^0} x_{\tau-1}^0$$

due to Lemma 2 (iv). Thus, we established the non-positiveness of each term of the summand for any period $\tau \geq 2$ in (48). Since, the latter is basically a finite weighted some of such terms with a negative sign in front, for any $a \in \mathbb{A}$, $\gamma_T^a \geq 0$. Finally we account for the coefficients multiplying $\bar{\theta}$ in equation (30) and show that

$$\begin{aligned} \mu_T &= \frac{\partial}{\partial \bar{\theta}} E \left[\sum_{\tau=2}^T \beta^{\tau-1} \left(-\alpha_1 (x_{\tau-1}^0 - x_\tau^0) \frac{\partial}{\partial x_1^0} (x_{\tau-1}^0 - x_\tau^0) + \alpha_2 (\theta_\tau^0 - x_\tau^0) \frac{\partial}{\partial x_1^0} x_\tau^0 \right. \right. \\ &\quad \left. \left. - \alpha_3 (x_\tau^{-1} - x_\tau^0) \frac{\partial}{\partial x_1^0} (x_\tau^{-1} - x_\tau^0) - \alpha_3 (x_\tau^1 - x_\tau^0) \frac{\partial}{\partial x_1^0} (x_\tau^1 - x_\tau^0) \right) \middle| (x^{t-1}, \theta^t) \right] \\ &\geq 0 \end{aligned} \quad (50)$$

Expectation washes out all individual θ_τ^a 's and we have only $\bar{\theta}$ apart from $(x_1^a)_{a \in \mathbb{A}}$ in each period's expression in (50). By symmetry of the form in Lemma (2) (i) across agents, the weight on $\bar{\theta}$ in x_t^0 , is

equal to the weight on $\bar{\theta}$ in x_t^1 and on x_t^{-1} . Thus, $\frac{\partial}{\partial \bar{\theta}} E[(x_t^0 - x_t^1)] = \frac{\partial}{\partial \bar{\theta}} E[(x_t^0 - x_t^{-1})] = 0$. This makes the second line of (50) equal to zero. By Lemma (2)-(i), the weight on $\bar{\theta}$ in $E[x_t^0]$, $1 - \Pi_{s=2}^t C_{T-(t-1)}$ (residual of the sum of the effects of $\{x_1^b\}$) is bigger than that in $E[x_{t-1}^0]$, $1 - \Pi_{s=2}^{t-1} C_{T-(t-1)}$; hence the term $\frac{\partial}{\partial \bar{\theta}} E[(x_{t-1}^0 - x_t^0)] \leq 0$. By Lemma (2)-(i), $\frac{\partial}{\partial \bar{\theta}} E[x_t^0] \geq 0$. By Lemma (2)-(iv), $\frac{\partial}{\partial x_1^0} (x_t^0 - x_{t-1}^0) \leq 0$. All these together imply that the expression in (50) is non-negative. Each $E[x_\tau^b]$ in (30) can be written as a convex combination of $(x_1^a)_{a \in \mathbb{A}}$, $\bar{\theta}$, x_0^0 , θ_1^0 , with the help of Lemma 2-(i). Since at each iteration, convex combination structure is preserved, it is so at the end too. Then, the sum of coefficients in each of the differences involving those variables in the parentheses is zero. This in turn implies that the total sum of coefficients in (30) is zero. Thus, the alternative formulation in (32) is true. This concludes the proof of Lemma 3. $\blacksquare$

Proof of Theorem 3 (Recursive Computation) : Consider a finite-horizon T -period economy with conformity preferences ($\alpha_i > 0$, $i = 1, 2, 3$) and complete information. For part (i), we simply assume that $T = 1$ and show that one can fit an exponentially declining sequence into equation 24. Since that equation has a unique solution as argued in the existence proof, that solution must have exponentially declining coefficients. Matching the coefficients of the policy function using equation 24, one gets for $a \neq 0$

$$d_1^{a+1} = \left(\frac{\alpha_3}{\Delta_1} \right) d_1^{a+2} + \left(\frac{\alpha_3}{\Delta_1} \right) d_1^a$$

Dividing both sides by d_1^a and multiplying them by $\left(\frac{\Delta_1}{\alpha_3} \right)$, one gets

$$\left(\frac{\Delta_1}{\alpha_3} \right) \underbrace{\left(\frac{d_1^{a+1}}{d_1^a} \right)}_{r_1} = \underbrace{\left(\frac{d_1^{a+2}}{d_1^{a+1}} \right)}_{r_1} \underbrace{\left(\frac{d_1^{a+1}}{d_1^a} \right)}_{r_1} + 1$$

which induces a quadratic equation

$$r_1^2 - \left(\frac{\Delta_1}{\alpha_3} \right) r_1 + 1 = 0$$

whose determinant $\left(\frac{\Delta_1}{\alpha_3} \right)^2 - 4 > 0$ since $\Delta_1 = \alpha_1 + \alpha_2 + 2\alpha_3 > 2\alpha_3$ (remember that $\alpha_i > 0$ for $i = 1, 2, 3$). The equation has two positive roots, one bigger and one smaller than 1. The bigger root cannot work since it is explosive as $|a| \rightarrow \infty$. We pick the smaller root

$$0 < r_1 = \left(\frac{\Delta_1}{2\alpha_3} \right) - \sqrt{\left(\frac{\Delta_1}{2\alpha_3} \right)^2 - 1} < 1 \quad (51)$$

which is decreasing in $\left(\frac{\Delta_1}{2\alpha_3} \right)$ spanning the interval $(0, 1)$ for different values of the former in the interval $(1, \infty)$. Finally, the sum of coefficients can be written

$$\begin{aligned} \sum_{a \in \mathbb{A}} d_1^a &= \sum_{a \in \mathbb{A}} d_1^0 r_1^{|a|} \\ &= d_1^0 + 2 d_1^0 \frac{r_1}{1 - r_1} \\ &= \frac{\alpha_2}{\alpha_1 + \alpha_2} \end{aligned} \quad (52)$$

The first equality is due to the exponentiality of the sequence; the third uses the same argument as in (47) with $\mu_1 = 0$, for the coefficient sequence $(d_1^a)_{a \in \mathbb{A}}$. Solving for d_1^0 from above, we obtain

$$d_1^0 = \left(\frac{\alpha_2}{\alpha_1 + \alpha_2} \right) \left(\frac{1 - r_1}{1 + r_1} \right)$$

and finally thanks to exponentiality

$$d_1^a = r_1^{|a|} \left(\frac{\alpha_2}{\alpha_1 + \alpha_2} \right) \left(\frac{1 - r_1}{1 + r_1} \right), \quad \text{for } a \in \mathbb{A}$$

The argument for the sequence $(c_1^a)_{a \in \mathbb{A}}$ is identical with one change: The sum of coefficients $\sum_a c_1^a = \left(\frac{\alpha_1}{\alpha_1 + \alpha_2} \right)$. This proves part (i) of the theorem.

For part (ii), observe that the parameters of the maps L_s , namely $\Delta_s, (\gamma_s^a), \mu_s$ are functions only of the continuation policy coefficients $(c_\tau^*, d_\tau^*, e_\tau^*)_{\tau=1}^{s-1}$ as defined in (31), (48), and (50), simply because these are “forward-looking” expressions. We saw in the induction step (**Step 2**) of the existence proof that L_s defined in this fashion becomes a contraction and has a unique fixed point, which is the coefficient sequence of the first-period policy of an s -period continuation. This establishes part (ii).

For part (iii), observe that each $g \in G$ is associated with coefficients $((c^a, d^a)_a, e)$. Clearly, for any sequence of policies in G , $g_n \rightarrow g$ in sup norm if and only if the associated coefficients $((c_n^a, d_n^a)_a, e_n) \rightarrow ((c^a, d^a)_a, e)$ in sup norm. In Step 3 of the existence proof, we establish the convergence of the finite-horizon equilibrium policies to the stationary infinite-horizon MPE policy as the horizon expands. But this implies that the associated unique coefficient sequence also should converge, then, to the coefficient sequence of the infinite-horizon stationary MPE policy. This establishes part (iii) of Theorem 3. ■

Proof of Proposition 1 (Tail Convergence Monotonicity) : The proof is by induction on T . For $T = 1$, we know from Theorem 3 (i) that the policy coefficient sequence $(d_1^a)_{a \in \mathbb{A}}$ is exponentially declining on both sides of the origin, at the rate r_1 . From the form of the policy function in Theorem 1, the conditional covariance between agents 0 and $a + 1$, with $a \geq 0$ w.l.o.g., given x_0 is

$$\begin{aligned} \text{Cov} \left(x_1^0, x_1^{a+1} \middle| x_0 \right) &= \text{Cov} \left(\sum_{b_1 \in \mathbb{A}} d_1^{b_1} \theta_1^{b_1}, \sum_{b_2 \in \mathbb{A}} d_1^{b_2} \theta_1^{a+1+b_2} \right) \\ &= \sum_{b_1 \in \mathbb{A}} d_1^{b_1} \text{Cov} \left(\theta_1^{b_1}, \sum_{b_2 \in \mathbb{A}} d_1^{b_2} \theta_1^{a+1+b_2} \right) \\ &= \text{Var}(\theta) \sum_{b_1 \in \mathbb{A}} d_1^{b_1} d_1^{b_1-(a+1)} \end{aligned} \tag{53}$$

We will focus on the summation term in the last expression in (53). Write it as

$$\begin{aligned} \sum_{b_1 \in \mathbb{A}} d_1^{b_1} d_1^{b_1-(a+1)} &= \sum_{b_1 < 0} d_1^{b_1} d_1^{b_1-(a+1)} + d_1^0 d_1^{-a-1} + \sum_{b_1 > 0} d_1^{b_1} d_1^{b_1-(a+1)} \\ &= \sum_{b_1 < 0} d_1^{b_1} d_1^{b_1-(a+1)} + d_1^0 d_1^{-a-1} + \sum_{b_1 \geq 0} d_1^{b_1+1} d_1^{b_1-(a)} \\ &= \sum_{b_1 < 0} d_1^{b_1} \left(r_1 d_1^{b_1-a} \right) + (d_1^0)^2 (r_1)^{a+1} + \sum_{b_1 \geq 0} \left(r_1 d_1^{b_1} \right) d_1^{b_1-a} \\ &= r_1 \sum_{b_1 \in \mathbb{A}} d_1^{b_1} d_1^{b_1-a} + (d_1^0)^2 (r_1)^{a+1} \\ &= r_1 \text{Var}(\theta)^{-1} \text{Cov} \left(x_1^0, x_1^a \middle| x_0 \right) + (d_1^0)^2 r_1^{a+1} \end{aligned}$$

The first equality is a partitioning, the second a simple change of variable, and the third is due to the symmetry and the exponentiality of the d_1^a sequence. Substituting the final expression back in (53) yields, for all $a \geq 0$

$$\text{Cov} \left(x_1^0, x_1^{a+1} \middle| x_0 \right) = r_1 \text{Cov} \left(x_1^0, x_1^a \middle| x_0 \right) + r_1^{a+1} \text{Var}(\theta) (d_1^0)^2 \quad (54)$$

which implies that the rate of decay of the covariances is greater than r_1 , for any $a \geq 0$. Since the second term on the right hand side of (54) decays at the rate r_1 , this implies that the ratio

$$\frac{r_1^{a+1} \text{Var}(\theta) (d_1^0)^2}{\text{Cov} \left(x_1^0, x_1^a \middle| x_0 \right)} \quad (55)$$

decreases monotonically, and being non-negative, it converges. Actually it converges to zero. Here is why. Since the ratio is less than 1, suppose that it converges to $k \in (0, 1)$. This means from (54) that the limit rate of decay of the covariances is $r_1 + k$, greater than the rate for the term in the numerator in (55). Thus, the ratio in (55) should converge to zero at the limit, a contradiction to $k \in (0, 1)$. So, the limit of (55) is zero, which in turn implies from (54), after dividing both sides by $\text{Cov} \left(x_1^0, x_1^a \middle| x_0 \right)$, that

$$\lim_{a \rightarrow \infty} \frac{\text{Cov} \left(x_1^0, x_1^{a+1} \middle| x_0 \right)}{\text{Cov} \left(x_1^0, x_1^a \middle| x_0 \right)} = r_1 \quad (56)$$

The argument is symmetric for $a \leq 0$; hence the sequence $\left\{ \text{Cov} \left(x_1^0, x_1^a \middle| x_0 \right) \right\}_{a \in \mathbb{A}}$ declines exponentially on both tails at the same rate r_1 and the statement is true for $T = 1$.

Now assume that the statement in Proposition 1 is true for economies up to $T - 1$ period. We will show that it should also hold for T -period economies. We will base the main induction arguments on the following Lemma.

Lemma 8 *The sequence $\{\gamma_T^b\}_{b \in \mathbb{A}}$ in Lemma 3 and the equilibrium coefficient sequence (c_T, d_T) for the first-period policy of a T -period economy have the following properties: The **rate** at which they decline **at the tail** satisfies, for $T \geq 2$*

$$\lim_{a \rightarrow \infty} \left(\frac{\gamma_T^{a+1}}{\gamma_T^a} \right) = \lim_{a \rightarrow -\infty} \left(\frac{\gamma_T^{a-1}}{\gamma_T^a} \right) = r_{T-1},$$

and

$$\lim_{a \rightarrow \infty} \left(\frac{d_T^{a+1}}{d_T^a} \right) = \lim_{a \rightarrow -\infty} \left(\frac{d_T^{a-1}}{d_T^a} \right) = r_T > r_{T-1}.$$

Proof of Lemma 8 : Let $u(t) := u(x_{t-1}^0, x_t^0, \{x_t^b\}_{b \in \{-1, 1\}}, \theta_t^0)$ where u represents the conformity preferences in Assumption 1. Let $u_0(t) := \frac{\partial}{\partial x_1^0} u(t)$. From equation (48), γ_T^a can be written as

$$\begin{aligned} \gamma_T^a &:= \alpha_3 I_{\{a \in \{-1, 1\}\}} \\ &+ \sum_{\tau=2}^T \beta^{\tau-1} \left[\left(\frac{\partial x_{\tau-1}^0}{\partial x_1^a} \right) \frac{\partial}{\partial x_{\tau-1}^0} u_0(\tau) + \left(\frac{\partial x_{\tau}^0}{\partial x_1^a} \right) \frac{\partial}{\partial x_{\tau}^0} u_0(\tau) + \left(\frac{\partial x_{\tau-1}^{-1}}{\partial x_1^a} \right) \frac{\partial}{\partial x_{\tau-1}^{-1}} u_0(\tau) + \left(\frac{\partial x_{\tau}^{-1}}{\partial x_1^a} \right) \frac{\partial}{\partial x_{\tau}^{-1}} u_0(\tau) \right] \end{aligned} \quad (57)$$

We will focus on the second term inside the summand since the method of proof will apply to the remaining terms straightforwardly. Assume w.l.o.g. that $a \geq 0$.

$$\begin{aligned} \left(\frac{\partial x_\tau^0}{\partial x_1^a} \right) \frac{\partial}{\partial x_\tau^0} u_0(\tau) &= \sum_{s \in \mathbb{A}} \left(\frac{\partial x_2^s}{\partial x_1^a} \right) \left(\frac{\partial x_\tau^0}{\partial x_2^s} \right) \frac{\partial}{\partial x_\tau^0} u_0(\tau) \\ &= \sum_{s \in \mathbb{A}} c_2^{a-s} \left(\frac{\partial x_\tau^0}{\partial x_2^s} \right) \frac{\partial}{\partial x_\tau^0} u_0(\tau) \end{aligned}$$

and the corresponding term for γ_T^{a+1} is

$$\sum_{s \in \mathbb{A}} c_2^{a+1-s} \left(\frac{\partial x_\tau^0}{\partial x_2^s} \right) \frac{\partial}{\partial x_\tau^0} u_0(\tau)$$

as $a \rightarrow \infty$ the tail convergence rate for the continuation kicks in and

$$\begin{aligned} \lim_{a \rightarrow \infty} \sum_{s \in \mathbb{A}} c_2^{a+1-s} \left(\frac{\partial x_\tau^0}{\partial x_2^s} \right) \frac{\partial}{\partial x_\tau^0} u(\tau) &= \lim_{a \rightarrow \infty} \sum_{s \in \mathbb{A}} r_{T-1} c_2^{a-s} \left(\frac{\partial x_\tau^0}{\partial x_2^s} \right) \frac{\partial}{\partial x_\tau^0} u(\tau) \\ &= r_{T-1} \lim_{a \rightarrow \infty} \sum_{s \in \mathbb{A}} c_2^{a-s} \left(\frac{\partial x_\tau^0}{\partial x_2^s} \right) \frac{\partial}{\partial x_\tau^0} u(\tau) \end{aligned}$$

thus

$$\lim_{a \rightarrow \infty} \left(\frac{\partial x_\tau^0}{\partial x_t^{a+1}} \right) \frac{\partial}{\partial x_\tau^0} u(\tau) = r_{T-1} \lim_{a \rightarrow \infty} \left(\frac{\partial x_\tau^0}{\partial x_1^a} \right) \frac{\partial}{\partial x_\tau^0} u(\tau)$$

Hence at the tail, the second term of the sum inside the brackets of 57 decays at the rate r_{T-1} . The same one-step transition argument applies to each terms of the sum, in equation (57). Moreover, since (57) is a discounted sum, the entire expression is summable and it implies that

$$\lim_{a \rightarrow \infty} \left(\frac{\gamma_T^{a+1}}{\gamma_T^a} \right) = r_{T-1}$$

inheriting the rate of tail convergence of the continuation economy, as is argued in the Lemma. The method of proof for $a \leq 0$ is identical thanks to the symmetry of the environment.

For the second part of the Lemma, let $D(r_{T-1})$ be the space of sequences that satisfies the properties in (25) and that converges at the tail at a rate $r_T \geq r_{T-1}$. This is a closed subset of the space of sequences that satisfy only the properties in (25), hence a complete metric space itself. Consequently, the unique coefficient sequence d_T that is the fixed point of the map in (33) should lie in $D(r_{T-1})$. Let $D'(r_{T-1}) \subset D(r_{T-1})$ be the space of sequences in $D(r_{T-1})$ whose convergence at the tail is *strictly greater* than r_{T-1} . We will show below that the map in (33) maps elements of $D(r_{T-1})$ into the set $D'(r_{T-1})$, which will imply that the unique solution of the map (33) converges at a rate $r_T > r_{T-1}$ at the tail.

Pick agent $2a + 1$ and assume w.l.o.g. that $a \geq 0$. Let $d_T \in D'(r_{T-1})$. From (33) by matching

coefficients

$$\begin{aligned}
\hat{d}_T^{2a+1} &= \Delta_T^{-1} \left[\sum_{b \neq 0} \gamma_T^b d_T^{2a+1-b} \right] \\
&= \Delta_T^{-1} \left[\sum_{b > a} \gamma_T^b d_T^{2a+1-b} + \gamma_T^a d_T^{a+1} + \sum_{b < a, b \neq 0} \gamma_T^b d_T^{2a+1-b} \right] \\
&= \Delta_T^{-1} \left[\sum_{b \geq a} \gamma_T^{b+1} d_T^{2a-b} + \gamma_T^a d_T^{a+1} + \sum_{b < a, b \neq 0} \gamma_T^b d_T^{2a+1-b} \right] \\
&= \Delta_T^{-1} \left[\sum_{b \geq a} \gamma_T^b d_T^{2a-b} \left(\frac{\gamma_T^{b+1}}{\gamma_T^b} \right) + \gamma_T^a d_T^{a+1} + \sum_{b < a, b \neq 0} \gamma_T^b d_T^{2a-b} \left(\frac{d_T^{2a+1-b}}{d_T^{2a-b}} \right) \right] \quad (58)
\end{aligned}$$

The second equality is a partitioning of the sum taking agent a as the ‘middle’; the first sum after the third equality is a simple shift and change of the dummy variable b ; the first term after the first equality sign is by multiplying and dividing each term in the summand by γ_T^b ; finally the last term after the fourth equality sign is by multiplying and dividing each term in the summand by d_T^{2a-b} . Since all elements involved are non-zero, the algebraic manipulation above is feasible. We can add to and subtract from equation (58) the term $\Delta_T^{-1} \sum_{b < a, b \neq 0} \gamma_T^b d_T^{2a-b} r_{T-1}$ and rearrange the order of the terms to obtain

$$\begin{aligned}
\hat{d}_T^{2a+1} &= \Delta_T^{-1} \left[\underbrace{\sum_{b \geq a} \gamma_T^b d_T^{2a-b} \left(\frac{\gamma_T^{b+1}}{\gamma_T^b} \right) + \sum_{b < a, b \neq 0} \gamma_T^b d_T^{2a-b} r_{T-1}}_A + \underbrace{\gamma_T^a d_T^{a+1}}_B \right. \\
&\quad \left. + \underbrace{\sum_{b < a, b \neq 0} \gamma_T^b d_T^{2a-b} \left\{ \left(\frac{d_T^{2a+1-b}}{d_T^{2a-b}} \right) - r_{T-1} \right\}}_C \right] \quad (59)
\end{aligned}$$

The analogous expression for $\hat{d}_T^{2a}$ is given, after a similar partitioning with agent a as the middle agent, by

$$\begin{aligned}
\hat{d}_T^{2a} &= \Delta_T^{-1} \left[\sum_{b \neq 0} \gamma_T^b d_T^{2a-b} \right] \\
&= \Delta_T^{-1} \left[\underbrace{\sum_{b \geq a} \gamma_T^b d_T^{2a-1-b} \left(\frac{\gamma_T^{b+1}}{\gamma_T^b} \right) + \sum_{b < a, b \neq 0} \gamma_T^b d_T^{2a-1-b} r_{T-1}}_{A'} + \underbrace{\gamma_T^a d_T^a}_{B'} \right. \\
&\quad \left. + \underbrace{\sum_{b < a, b \neq 0} \gamma_T^b d_T^{2a-1-b} \left\{ \left(\frac{d_T^{2a-b}}{d_T^{2a-1-b}} \right) - r_{T-1} \right\}}_{C'} \right] \quad (60)
\end{aligned}$$

We showed above in the first part of the proof of Lemma 8 that as $b \rightarrow \infty$, the ratio $(\gamma_T^{b+1}/\gamma_T^b) \rightarrow r_{T-1}$. Consequently, the following limits hold

$$\lim_{a \rightarrow \infty} \Delta_T^{-1} A = r_{T-1} \lim_{a \rightarrow \infty} \hat{d}_T^{2a} \quad (61)$$

$$\lim_{a \rightarrow \infty} \Delta_T^{-1} A' = r_{T-1} \lim_{a \rightarrow \infty} \hat{d}_T^{2a-1} \quad (62)$$

$$\lim_{a \rightarrow \infty} \Delta_T^{-1} C = (r_T - r_{T-1}) \lim_{a \rightarrow \infty} \hat{d}_T^{2a} \quad (63)$$

$$\lim_{a \rightarrow \infty} \Delta_T^{-1} C' = (r_T - r_{T-1}) \lim_{a \rightarrow \infty} \hat{d}_T^{2a-1} \quad (64)$$

The expressions in (61) and (63) put together imply that as a gets large the ratio

$$\left(\frac{\hat{d}_T^{2a+1}}{\hat{d}_T^{2a}} \right) \approx r_T + \frac{\gamma_T^a d_T^{a+1}}{\hat{d}_T^{2a}} \geq r_T \quad (65)$$

and the expressions in (62) and (64) put together imply that

$$\left(\frac{\hat{d}_T^{2a}}{\hat{d}_T^{2a-1}} \right) \approx r_T + \frac{\gamma_T^a d_T^a}{\hat{d}_T^{2a-1}} \geq r_T \quad (66)$$

The last two expressions imply that the ratios $\frac{\gamma_T^a d_T^{a+1}}{\hat{d}_T^{2a}}$ and $\frac{\gamma_T^a d_T^a}{\hat{d}_T^{2a-1}}$ converge. This is because as a gets arbitrarily large, the numerator converges at the rate $r_{T-1} r_T$ and the denominator at a rate greater than equal to r_T^2 . Since both ratios are strictly less than one, one of the following two possibilities must hold: either (i) they converge to a positive constant less than one (the case where $r_T = r_{T-1}$) or (ii) they converge to zero (the case where $r_T > r_{T-1}$). The first case is not possible. Suppose it is. Then, $r_T = r_{T-1}$. This implies from (59) along with (61) and (63) that

$$r_T = \lim_{a \rightarrow \infty} \left(\frac{\hat{d}_T^{2a+1}}{\hat{d}_T^{2a}} \right) = r_T + \lim_{a \rightarrow \infty} \left(\frac{\gamma_T^a d_T^{a+1}}{\hat{d}_T^{2a}} \right) > r_T \quad (67)$$

a contradiction. Therefore the second case (ii) must be true. The argument for $a \leq 0$ is symmetric. This means that any sequence $d_T \in D(r_{T-1})$ is mapped to a sequence $\hat{d}_T \in D'(r_{T-1})$, meaning that it converges at the rate $r_T > r_{T-1}$ at the tail. So our claim in the beginning is true and the unique sequence d_T that satisfies properties in (25) and the equation (33) converges at the tail at a rate $r_T > r_{T-1}$. This concludes the proof of Lemma 8. $\blacksquare$

Rest of the ‘Proof of Proposition 1’ Now assume that the statement in Proposition 1 is true for economies up to $T - 1$ period. We will show that it should also hold for T -period economies. Consider

first the covariance between agents 0 and $2a + 1$, with $a \geq 0$ w.l.o.g.

$$\begin{aligned}
Cov\left(x_1^0, x_1^{2a+1} \middle| x_0\right) &= Var(\theta) \sum_{b \in \mathbb{A}} d_1^b d_1^{b-(2a+1)} \\
&= Var(\theta) \left[\sum_{b \leq a} d_T^b d_T^{b-(2a+1)} + d_T^{a+1} d_T^{-a} + \sum_{b \geq a+2} d_T^b d_T^{b-(2a+1)} \right] \\
&= Var(\theta) \left[\sum_{b \leq a} d_T^b d_T^{b-(2a+1)} + d_T^{a+1} d_T^{-a} + \sum_{b \geq a+1} d_T^{b+1} d_T^{b+1-(2a+1)} \right] \\
&= Var(\theta) \left[\sum_{b \leq a} d_T^b d_T^{b-(2a)} \left(\frac{d_T^{b-(2a+1)}}{d_T^{b-(2a)}} \right) + \sum_{b \geq a+1} d_T^b d_T^{b-(2a)} \left(\frac{d_T^{b+1}}{d_T^b} \right) + d_T^{a+1} d_T^{-a} \right]
\end{aligned}$$

where the algebraic manipulation is the same as in the proof of Lemma 8. The analogous expression for agent $2a$, taking agent a as the agent in the middle, is

$$Cov\left(x_1^0, x_1^{2a} \middle| x_0\right) = Var(\theta) \left[\sum_{b \leq a} d_T^b d_T^{b-(2a-1)} \left(\frac{d_T^{b-2a}}{d_T^{b-(2a-1)}} \right) + \sum_{b \geq a+1} d_T^b d_T^{b-(2a-1)} \left(\frac{d_T^{b+1}}{d_T^b} \right) + d_T^a d_T^{-a} \right]$$

We know from Lemma 8 that as $a \rightarrow \infty$, the ratio $(d_T^{a+1}/d_T^a) \rightarrow r_T > r_{T-1}$. This implies, from the expressions above for the covariance terms, that for large a ,

$$\frac{Cov\left(x_1^0, x_1^{2a+1} \middle| x_0\right)}{Cov\left(x_1^0, x_1^{2a} \middle| x_0\right)} \approx r_T + \frac{d_T^{a+1} d_T^{-a}}{Cov\left(x_1^0, x_1^{2a} \middle| x_0\right)} \geq r_T > r_{T-1} \quad (68)$$

and

$$\frac{Cov\left(x_1^0, x_1^{2a} \middle| x_0\right)}{Cov\left(x_1^0, x_1^{2a-1} \middle| x_0\right)} \approx r_T + \frac{d_T^a d_T^{-a}}{Cov\left(x_1^0, x_1^{2a-1} \middle| x_0\right)} \geq r_T > r_{T-1} \quad (69)$$

and straightforward modifications of the argument used in the proof of Lemma 8 implies that the ratios $d_T^{a+1} d_T^{-a} Cov\left(x_1^0, x_1^{2a} \middle| x_0\right)^{-1}$ and $d_T^a d_T^{-a} Cov\left(x_1^0, x_1^{2a-1} \middle| x_0\right)^{-1}$ both converge to zero and one obtains

$$\lim_{a \rightarrow \infty} \frac{Cov\left(x_1^0, x_1^{2a+1} \middle| x_0\right)}{Cov\left(x_1^0, x_1^{2a} \middle| x_0\right)} = r_T > r_{T-1} \quad (70)$$

$$\lim_{a \rightarrow \infty} \frac{Cov\left(x_1^0, x_1^{2a} \middle| x_0\right)}{Cov\left(x_1^0, x_1^{2a-1} \middle| x_0\right)} = r_T > r_{T-1} \quad (71)$$

thus the statement of Proposition 1 is true for any finite T -period economy. Clearly, $r_T \leq 1$ for any $T \geq 1$ since the non-negative d sequences sum up to less than 1. Hence, what we have is a monotone increasing sequence bounded from above by 1. Hence, the limit $r_\infty = \lim_{T \rightarrow \infty} r_T$ exists and is less than or equal to 1. Moreover, we know from Theorem 3 that the sequence of finite-horizon MPE coefficients converges

to that of the infinite-horizon MPE coefficient sequence d , thus r_∞ is the tail convergence rate of the infinite-horizon MPE coefficient sequence d . Therefore $r_\infty < 1$ since otherwise that would contradict the summability of the sequence d . This establishes the proof of Proposition 1. $\blacksquare$

Proof of Proposition 2 We showed in the proof of Proposition 1 that, for $T = 1$, the ratio $\left(\frac{\rho_{a+1,T}}{\rho_{a,T}}\right)$ is necessarily monotonically decreasing in a for any underlying preference parameter vector α , converging eventually, at the tail, to the rate r_1 given in Theorem 3 (i). As we showed in Section 4.2, the cross-sectional covariances at the stationary distribution can be written recursively given the weights of the policy function. For the myopic policy function, they take the form

$$\text{Cov}(x^0, x^a) = \sum_{a_1 \in \mathbb{A}} \sum_{b_1 \in \mathbb{A}} c_1^{a_1} c_1^{b_1} \text{Cov}(x^{a_1}, x^{a+b_1}) + \text{Var}(\theta) \sum_{a_1 \in \mathbb{A}} d_1^{a_1} d_1^{a-a_1}, \quad (72)$$

Since the c_1 and d_1 sequences are exponential at the rate r_1 from Theorem 3 (i), by straightforward modifications of the arguments in the first part of the proof of Proposition 1, one can show that the ratio of consecutive covariances for the myopic, $\left(\frac{\text{Cov}(x^0, x^{a+1})}{\text{Cov}(x^0, x^a)}\right)$ converges monotonically for $a \geq 0$ as gets large.

As we presented in Figure 11, however, the above ratio for the stationary policy function is non-monotonic for a set of parameter values. Moreover, the map in (5) that generates the policy weights as fixed points is continuous in the parameters $(\alpha_1, \alpha_2, \alpha_3)$ of the utility function. Thus, for each element of the above set of parameters for which the ratio of consecutive stationary covariances converges non-monotonically, there is an open-neighborhood around it such that for each element $\hat{\alpha}$ of the neighborhood, the same non-monotonicity property obtains. This concludes the proof. $\blacksquare$

Proof of Lemma 4 (Compactness of L_β and G) Let $(\beta_{T_n})_n$ be a sequence lying in L_β that converges to $x = (x_t) \in [0, 1]^\infty$. This implies that $\beta_{T_n, t} \rightarrow x_t$, for all $t \geq 1$, which in turn means that $x_t \in \{0, \beta^t\}$ by the construction of L_β . Moreover, if $x_t = 0$ for some t , $x_{t+\tau} = 0$ for all $\tau \geq 1$ since the terms β_{T_n} are geometric (finite or infinite) sequences. There are two possibilities: either $x = (1, \beta, \dots, \beta^T, 0, 0, \dots)$ or $x = \beta^t$ for all $t \geq 1$. Both lie in L_β which means that the limit of any convergent sequence in L_β lies in L_β . This establishes that L_β is closed. Given any $\epsilon > 0$, choose $N \geq 1$, a natural number, s.t. $\beta^N < \epsilon$. It is easy to see that any element in L_β lies in the ϵ -neighborhood (with respect to the sup metric) of one of the elements in the finite set $\{\beta_1, \beta_2, \dots, \beta_N\} \subset L_\beta$. This establishes that L_β is totally bounded. Therefore, L_β is compact. We next show that G endowed with the sup norm is compact.

Let $H := \{x = (x^a)_{a \in \mathbb{A}} \mid x^a \leq (\frac{1}{2a}), \text{ for all } a \in \mathbb{A}\}$. Defined by inequality constraints, this set is closed under the sup norm. We will show that it is also totally bounded. For a given $\epsilon > 0$, one can find an $N \geq 1$ s.t. $\frac{1}{2N} < \epsilon$. Pick a sequence $\bar{x} \in H$. For any $a \in \mathbb{A}$ s.t. $|a| \geq N$, $[0, (2N)^{-1}] \subset B_\infty(x^a, \epsilon)$, the ϵ -ball around x^a with respect to the sup norm. For $|a| \leq N$, let $Y(a) := \{0, \epsilon, 2\epsilon, \dots, k_a \epsilon, (2a)^{-1}\}$, where k_a is the greatest integer s.t. $k_a \epsilon \leq (2a)^{-1}$. The set

$$\left\{ x \in H \mid x^a = \bar{x}^a, \text{ for } |a| \geq N, \text{ and } (x^{-(N-1)}, \dots, x^0, \dots, x^{N-1}) \in \prod_{|a| \leq N} Y(a), \text{ for } |a| \leq N \right\}$$

is a finite set of elements of H . Moreover, it is dense in H by construction. This establishes that H is totally bounded. Thus, H is compact under the sup norm.

Each $g \in G$ is associated with coefficients $((c^a, d^a)_a, e)$. Clearly, for any sequence of policies in G , $g_n \rightarrow g$ in sup norm if and only if the associated coefficients $((c_n^a, d_n^a)_a, e_n) \rightarrow ((c^a, d^a)_a, e)$ in sup norm.

We know from (25) that c satisfies properties (i), (ii) and (iii). Thus, for any $a \in \mathbb{A}$, $c^0 > c^1 > \dots > c^{|a|}$, $c^a = c^{-a}$ and $\sum_{|b| \leq |a|} c^b < 1$. Combining all these, we have $2|a|c^a < \sum_{|b| \leq |a|} c^b < 1$ which in turn implies that $c^a < \frac{1}{2|a|}$, for all $a \in \mathbb{A}$. Same bounds hold for the d sequence. But then, the space of associated coefficient sequences, call it L_G , can be seen as a closed subset of H , a compact metric. Consequently, L_G is compact, thus sequentially compact. Pick a sequence $(g_n) \in G$ and let (c_n, d_n, e_n) be the associated coefficient sequence lying in L_G . Since L_G is sequentially compact, there exists a subsequence $(c_{m_n}, d_{m_n}, e_{m_n}) \rightarrow (c, d, e) \in L_G$. The latter, being an admissible coefficient sequence, is associated with the policy $g(x, \theta) := \sum_a c^a x^a + \sum_a d^a \theta^a + e\bar{\theta}$. Thus, the respective policy subsequence $g_{m_n} \rightarrow g \in G$. This establishes that G is sequentially compact hence compact. This concludes Lemma 4. $\blacksquare$

Proof of Lemma 5 (Continuity): Since G endowed with the sup norm is a compact metric space due to Lemma 4, the metric $d(g, g') := \sum_{t=1}^{\infty} 2^{-t} \|g_t - g'_t\|_{\infty}$ induces the product topology on G^{∞} (see e.g., Aliprantis and Border (2006, p. 90)), where $\|\cdot\|_{\infty}$ is the supnorm as before. Let $(\beta_T, g) \in L_{\beta} \times G^{\infty}$ and $\epsilon > 0$ be given. Set $\epsilon' := (\frac{1-\beta}{1-\beta^{T+1}}) \epsilon$. The period utility u is uniformly continuous since X is compact. Thus, one can choose a $\delta' > 0$ such that for any t , $|x_t^0 - y_t^0| < \delta'$ implies

$$|u(x_{t-1}^0, x_t^0, \{x_t^b(g)\}_{b \in N(0)}, \theta_t^0) - u(y_{t-1}^0, y_t^0, \{x_t^b(g)\}_{b \in N(0)}, \theta_t^0)| < \epsilon'.$$

Set $\delta = 2^{-T} \delta'$. Pick $g^0, g'^0 \in \Gamma(\beta^T, g)$ such that $d(g^0, g'^0) < \delta$. This implies that for all $t \leq T$, $\|g_t^0 - g_t'^0\|_{\infty} < 2^T \delta = \delta'$ hence $|x_t^0(g^0) - x_t^0(g'^0)| < \delta$. Uniform continuity of u then implies that the period utility levels are uniformly bounded above by ϵ' for all periods $t \leq T$. The claim therefore follows from

$$|U(g^0; \beta_T, g) - U(g'^0; \beta_T, g)| < \frac{1 - \beta^{T+1}}{1 - \beta} \epsilon' = \epsilon$$

$\blacksquare$

Proof of Lemma 6 (Planner's First Order Condition): The proof uses an extension of the usual calculus of variation technique to our symmetric strategic environment. We prove it for the class of bounded, continuous, and measurable, real-valued functions on $\mathbf{X} \times \Theta$. Then, we use the restriction of the result to a subset of it, the space of bounded, continuous, and measurable, X -valued functions. Suppose that the function h provides the maximum for the planner's problem. For any other admissible function h' , define $k = h' - h$. Consider now the expected utility from a one-parameter deviation from the optimal policy h , i.e.,

$$J(a) := \int u(x_{T-1}^0, (h + ak)(x_{T-1}, \theta_T), (h + ak)(R^{-1} x_{T-1}, R^{-1} \theta_T), \quad (73)$$

$$(h + ak)(R x_{T-1}, R \theta_T), \theta_T^0) \mathbb{P}(d\theta_T) \pi_{T-1}(dx_{T-1}) \quad (74)$$

where a is an arbitrary real number and u represents the conformity preferences in Assumption 1.. Since h maximizes the planner's problem, the function J must assume its maximum at $a = 0$. Leibnitz's rule for differentiation under an integral along with the chain rule for differentiation gives us

$$J'(a) := \int (u_2 k + u_3 k \circ R^{-1} + u_4 k \circ R) d\mathbb{P} d\pi_{T-1}$$

where u_i is the partial derivative of u with respect to the i -th argument. For J to assume its maximum at $a = 0$, it must satisfy

$$\begin{aligned}
J'(0) &:= \int \left[u_2(x_{T-1}^0, h(x_{T-1}, \theta_T), h(R^{-1}x_{T-1}, R^{-1}\theta_T), h(Rx_{T-1}, R\theta_T), \theta_T^0) k(x_{T-1}, \theta_T) \right. \\
&+ u_3(x_{T-1}^0, h(x_{T-1}, \theta_T), h(R^{-1}x_{T-1}, R^{-1}\theta_T), h(Rx_{T-1}, R\theta_T), \theta_T^0) k(R^{-1}x_{T-1}, R^{-1}\theta_T) \\
&+ \left. u_4(x_{T-1}^0, h(x_{T-1}, \theta_T), h(R^{-1}x_{T-1}, R^{-1}\theta_T), h(Rx_{T-1}, R\theta_T), \theta_T^0) k(Rx_{T-1}, R\theta_T) \right] \\
&\times \mathbb{P}(d\theta_T) \pi_{T-1}(dx_{T-1}) = 0
\end{aligned} \tag{75}$$

for any arbitrary admissible deviation k . Suppose that the statement of the lemma is not true. This implies that there is an element $(\bar{x}, \bar{\theta}) \in \mathbf{X} \times \Theta$ such that

$$\begin{aligned}
0 &\neq u_2(\bar{x}^0, h(\bar{x}, \bar{\theta}), h(R^{-1}\bar{x}, R^{-1}\bar{\theta}), h(R\bar{x}, R\bar{\theta}), \bar{\theta}^0) \\
&+ u_3(\bar{x}^1, h(R\bar{x}, R\bar{\theta}), h(\bar{x}, \bar{\theta}), h(R^2\bar{x}, R^2\bar{\theta}), \bar{\theta}^1) \\
&+ u_4(\bar{x}^{-1}, h(R^{-1}\bar{x}, R^{-1}\bar{\theta}), h(R^{-2}\bar{x}, R^{-2}\bar{\theta}), h(\bar{x}, \bar{\theta}), \bar{\theta}^{-1})
\end{aligned} \tag{76}$$

Assume w.l.o.g. that the above expression takes a positive value (the proof for the case with a negative value is identical). Since the utility function, its partials, and the deviation functions are all continuous with respect to the product topology, and that the measures π and $\mathbb{P}$ have positive densities, there exists a $(\pi \times \mathbb{P})$ -positive measure neighborhood $A \subset \mathbf{X} \times \Theta$ around $(\bar{x}, \bar{\theta})$ such that the above expression stays positive for all $(x_{T-1}, \theta_T) \in A$.⁷⁰ Assume that $a_1 = (\bar{x}, \bar{\theta})$, $a_2 = (R\bar{x}, R\bar{\theta})$, and $a_3 = (R^{-1}\bar{x}, R^{-1}\bar{\theta})$ are distinct points. Otherwise, since the underlying space X is a real interval and the maps R and R^{-1} are right and left shift maps, one can always pick a point in A that has that property.

Now choose $\epsilon > 0$ small enough so that the ϵ -balls $B_\epsilon(a_1)$, $B_\epsilon(a_2)$, and $B_\epsilon(a_3)$ are disjoint. R and R^{-1} being both continuous are homeomorphisms. So, one can find $\epsilon > \delta_1 > 0$ and $\epsilon > \delta_2 > 0$ such that $R(B_{\delta_1}(a_1)) \subset B_\epsilon(a_2)$ and $R^{-1}(B_{\delta_2}(a_1)) \subset B_\epsilon(a_3)$. Let $\delta = \min\{\delta_1, \delta_2\}$ and $A_1 := B_\delta(a_1)$. We next define a particular deviation k . Let the function k be defined as

$$k(x, \theta) = k(Rx, R\theta) = k(R^{-1}x, R^{-1}\theta) = \begin{cases} \gamma [\delta - d((x, \theta), a_1)], & \text{if } (x, \theta) \in A_1 \\ 0, & \text{otherwise.} \end{cases} \tag{77}$$

where $\gamma > 0$ is a scalable constant. This is possible because A_1 , $R(A_1)$ and $R^{-1}(A_1)$ are disjoint sets. Constructed this way, k is a bounded, continuous, and measurable function⁷¹. Substitute k into equation (75). By construction, the only set on which k is positive is the set A_1 which is itself a subset of A , the set of elements of $\mathbf{X} \times \Theta$ for which the expression (76) is positive. Hence, evaluated with the constructed deviation function k , $J'(0) > 0$, a contradiction to the fact that the policy function h was optimal. Therefore the statement of the lemma must be true. This concludes the proof. $\blacksquare$

⁷⁰Endowed with the product topology, the space $\mathbf{X} \times \Theta$ is metrizable by the metric d . See footnote 65. Product topology and the associated metric allows us to choose positive measure proper subsets of X for choices of near-by agents and the whole sets X and Θ for far-away agents, staying at the same time in the close vicinity of the point $(\bar{x}, \bar{\theta})$.

⁷¹We endow the range space, the real line, with the Borel σ -field hence any continuous function into the real line is automatically measurable.

Proof of Lemma 7 Using (39) iteratively, one can write for any $(x_T, \theta_{T+1}) \in \mathbf{X} \times \Theta$

$$\begin{aligned} V_a^h(x_T, \theta_{T+1}) &= \int \sum_{t=T+1}^{T+N} \beta^{t-T-1} \left[2\alpha_1 (x_t^0 - x_{t-1}^0) \frac{\partial}{\partial x_T^a} (x_{t-1}^0 - x_t^0) + 2\alpha_2 (\theta_t^0 - x_t^0) \frac{\partial}{\partial x_T^a} x_t^0 \right. \\ &\quad + 2\alpha_3 (x_t^{-1} - x_t^0) \frac{\partial}{\partial x_T^a} (x_t^{-1} - x_t^0) + 2\alpha_3 (x_t^1 - x_t^0) \frac{\partial}{\partial x_T^a} (x_t^1 - x_t^0) \\ &\quad \left. + \beta^{N+1} V_a^h(x_{t+N}, \theta_{t+N+1}) \right] \prod_{i=1}^N \mathbb{P}(d\theta_{T+1+i}) \end{aligned} \quad (78)$$

where x_t is written as, using iterations of the policy function g and Lemma 2 (i) with x_T instead of x_1

$$\begin{aligned} x_t^a &= \sum_{b_1 \in \mathbb{A}} \dots \sum_{b_{t-T} \in \mathbb{A}} c^{b_1} \dots c^{b_{t-T}} x_T^{a+b_1+\dots+b_{t-T}} \\ &\quad + \sum_{s=1}^{t-T} \sum_{b_1 \in \mathbb{A}} \dots \sum_{b_{s-1} \in \mathbb{A}} c^{b_1} \dots c^{b_{s-1}} \left(\sum_{b_s \in \mathbb{A}} d^{b_s} \theta_{t-(s-1)}^{a+b_1+\dots+b_s} + e^{\bar{\theta}} \right) \end{aligned} \quad (79)$$

At the point $(\hat{x}, \hat{\theta})$, $x_T^a = \bar{x}$ for all $a \in \mathbb{A}$. So, the first part after the equality sign in (79) is the same for all agents. Since the preference shocks are i.i.d., the second part will be the same for all agents in expectations, which eliminates the terms in the second line after the equality sign in (78). Thanks to Lemma 2 (i), $\frac{\partial}{\partial x_T^a} x_t^0 > 0$ for any $a \in \mathbb{A}$, and for all $t = T+1, \dots, T+N$. But then, the second term in (78) after the first bracket is negative in expectations. This is because using (79) $E[x_t^0 | (x_T, \theta_{T+1})] = C^{t-T} \bar{x} + (1 - C^{t-T}) \bar{\theta} > \bar{\theta}$, where $C = \sum_a c^a$. The first term after the bracket sign too is negative in expectations. Here is why: The term

$$\begin{aligned} E[(x_t^0 - x_{t-1}^0) | (x_T, \theta_{T+1})] &= C^{t-T} \bar{x} + (1 - C^{t-T}) \bar{\theta} - C^{t-1-T} \bar{x} - (1 - C^{t-1-T}) \bar{\theta} \\ &= C^{t-1-T} (1 - C) (\bar{\theta} - \bar{x}) < 0 \end{aligned}$$

for any $t = T+1, \dots, T+N$. So, one can write

$$E[2\alpha_1 (x_t^0 - x_{t-1}^0) \frac{\partial}{\partial x_T^a} (x_{t-1}^0 - x_t^0) | (x_T, \theta_{T+1})] < E[2\alpha_1 (x_t^0 - x_{t-1}^0) \frac{\partial}{\partial x_T^a} x_{t-1}^0 | (x_T, \theta_{T+1})] < 0$$

which shows that the summand in (78) is negative in expectations in every period. In turn, the whole sum, then, until the last line of (78), is negative in expectations for any arbitrary N . The choice of a was arbitrary and that V_a^h is continuous on $\mathbf{X} \times \Theta$ for any $a \in \mathbb{A}$. The latter is compact with respect to the product topology. Hence, V_a^h is bounded. So, one can choose an N large enough to make the $\beta^{N+1} V_a^h(x_{t+N}, \theta_{t+N+1})$ term arbitrarily small. This implies that the whole expression in (78) is negative, which in turn means that $V_a^h(\hat{x}, \hat{\theta}) < 0$ for any $a \in \mathbb{A}$.

At the point $(\hat{x}, \hat{\theta})$, the first line of (40) is zero and the second line is negative, as we just showed, which makes the whole expression in (40) negative. Since the first line in (40) is continuous and so are V_a for any $a \in \mathbb{A}$, the whole expression in (40) is continuous. Hence, as in the proof of Lemma 6, there exists a $(\pi \times \mathbb{P})$ -positive measure neighborhood $E \subset \mathbf{X} \times \Theta$ around $(\bar{x}, \bar{\theta})$ such that the above expression stays negative for all $(x_{T-1}, \theta_T) \in E$. This concludes the proof. $\blacksquare$

14 Appendix E: Details about the simulations

We build an artificial economy that consists of a large number of agents ($|\mathbb{A}| = 1300, 2500$, and 5000 , depending on the treatment) distributed on the one-dimensional integer lattice. At both ends “buffer” agents that act randomly are added to smooth boundary effects. Depending on the treatment, we start the economy with the following initial configuration of choices: (i) the highest action for all agents; (ii) the lowest action for all agents, (iii) the action equal to the mean shock for all agents.

The core engine behind the simulations is a Matlab code, **g.m**, which computes the equilibrium policy weights recursively as outlined in Section 3.2 of the paper. The code is posted on Özgür’s webpage, <http://www.sceco.umontreal.ca/onurozgur/>, at the Université de Montréal; the code contains also detailed explanations. The correlation computations use another code, **cor.m**, also available on Özgür’s webpage.

Both codes use as input parameters values of the preference parameters α_i , $i = 1, 2, 3$, the discount factor β , the horizon for the economy T , the number of agents $|\mathbb{A}|$, and the longest distance between agents for which the equilibrium correlation is computed M .

For the limit distributions results, once **g.m** computes the policy weights, we let the computer draw $(\theta_t^a)_{a=1}^{|\mathbb{A}|}$ from the interval $[-D, D]$ according to the uniform distribution (this is for simplicity since all results in the paper are distribution-free).