Dynamic Linear Economies with Social Interactions*

Onur Özgür[†]
Melbourne Business School
onur.ozgur@mbs.edu

Alberto Bisin[‡]
New York University
alberto.bisin@nyu.edu

Yann Bramoullé[§]
Aix-Marseille University, CNRS
yann.bramoulle@univ-amu.fr

May 22, 2018

Abstract

Social interactions are arguably at the root of several important socio-economic phenomena, from smoking and other risky behavioural patterns in teens to peer effects in school performance. We study *social interactions* in linear *dynamic* economies. For these economies, we are able to (i) obtain several desirable theoretical properties, such as existence, uniqueness, ergodicity; to (ii) develop simple recursive methods to rapidly compute equilibria; and to (iii) characterize several general properties of dynamic equilibria. Furthermore, we show that dynamic forward looking behaviour at equilibrium plays an instrumental role in allowing us to (iv) prove a positive identification result both in stationary and non-stationary economies. Finally, we study and sign the bias associated to disregarding dynamic equilibrium, e.g., postulating a sequence of static (myopic) one-period economies, a common practice in empirical work.

Journal of Economic Literature Classification Numbers: C18, C33, C62, C63, C73.
Keywords: Social interactions, conformity, ergodicity, habits, identification, Markov perfect equilibrium.

---

*We are grateful to Steven Durlauf, Matt Jackson, David Levine and Bernard Salanié, as well as to Victor Aguirregabiria, Ben Golub, Itai Sher, Boğaçhan Çelen, and Giulio Zanella for their extensive comments and suggestions at different stages. We thank Stéphane Bonhomme and four referees for their very careful and thoughtful editorial work on a previous draft. We also thank Massimiliano Amarante, Olivier Armantier, Jess Benhabib, Ken Binmore, Larry Blume, Michele Boldrin, Pierre-André Chiappori, In-Koo Cho, Catherine de Fontenay, Bryan Graham, Ali Hortaçsu, Frédéric Koessler, Takashi Kunimoto, Bob Lucas, Sophie Monnier, Áureo De Paula, Debraj Ray, Manuel Santos, Tom Sargent, José Scheinkman, Karl Schlag, Bruno Strulovici, Adam Szeidl, Jean Tirole, and many seminar participants. 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, SSHRC, FQRSC, IFM2, and Melbourne Business School.

[†]Melbourne Business School, 200 Leicester Street, Carlton, VIC 3053, Australia; CIREQ, CIRANO.

[‡]Department of Economics, New York University, 19 West Fourth Street, New York, NY, 10012, USA; NBER, CEPR.

[§]Aix-Marseille University, Aix-Marseille School of Economics, 5 Boulevard Bourdet, 13001 Marseille, France; CNRS, QREQAM.

1 Introduction

Agents interact in markets as well as socially, 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 interactions with others, by their positions in a predetermined network of relationships, e.g., a family, a peer group, or more generally any socioeconomic group. Social interactions are arguably at the root of several important phenomena. Peer effects, in particular, have been indicated as one of the main empirical determinants of risky behaviour in adolescents. Relatedly, peer effects have been studied in connection with education outcomes, obesity, friendship and sex, as well as in labor market referrals, neighborhood and employment segregation, criminal activity, and several other socioeconomic phenomena.[1]

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 act myopically. The theoretical and empirical study of dynamic economies with social interactions has in fact been hindered by several obstacles. Theoretically, the analysis of equilibria 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 with their large state space. Finally, in applications and empirical work, social interactions are typically hardly identified, even with population data.

In this paper, however, we study dynamic economies with social interactions, showing how some of these obstacles can be overcome. We are motivated by the fact that, in most applications of interest, social interactions are affected or constrained by relevant state variables. Indeed, peer effects act differently on individuals in different (non freely-reversible) states: belonging to a social group whose members are actively engaging in criminal activities affect agents with and without previous criminal experience differently; social links with female peers with an active job market occupation has a different effect on young women's labor market entry, exit, and career path decisions depending on their mothers' employment history; and so on. Furthermore, several forms of risky behaviour among adolescents induced by social interactions involve substance abuse and hence (the fundamentally dynamic) issues of addiction and habits. Indeed, dynamic equilibrium considerations have fundamental effects on the properties of economies with social interactions. In a dynamic equilibrium of a finite-horizon economy, for instance, agents whose choice is affected by the choices of their peers will rationally anticipate the expected length of

[1] See Brock and Durlauf (2001b), Glaeser and Scheinkman (2001), Glaeser and Scheinkman (2003), Moody (2001) for surveys; see also the *Handbook of Social Economics*, Benhabib, Bisin, and Jackson (2011).

these interactions, which in turn will affect their propensity for conformity.

We focus our attention on *linear economies*, in which each agent's preferences display preferences for *conformity*, that is, preferences which incorporate the desire to conform to the choices of agents in a reference group. More specifically, in our economy, each agent's instantaneous preferences depend on random preference shocks and on the current choices of agents in his social reference group, as a direct externality. We focus on symmetric structures of social interactions in reference groups, with agents disposed on the line or on the circle. Each agent's instantaneous preferences also depend on the agent's own previous choice, representing the inherent costs of dynamic behavioural changes due, e.g., to irreversibility, habits, or both. We do not impose any substantial restriction ex-ante on the random preference shocks hitting agents over time. In particular, specific forms of correlation across agents and time could capture selection in the formation of the social reference groups agents interact within.

Agents' choices at equilibrium are determined by *linear* policy (best reply) functions and depend on the previous choices and current preference shocks of all the other agents in the economy, as long as they are observable. In the special case of infinite-horizon economies with non-autocorrelated i.i.d. preference shocks and agents disposed on the line $\mathbb{A}$, a symmetric Markov perfect equilibrium (MPE) is represented by a symmetric policy function, g, which maps an agent's current choice at time t, linearly in each agent's past choices, $y_{b,t-1}$, in each agent's contemporaneous idiosyncratic preference shock, $\theta_{b,t}$, and in its mean, $E(\theta)$:

$$g\left(y_{t-1}, \theta_t\right) = \sum_{b \in \mathbb{A}} c_b \, y_{b,t-1} + \sum_{b \in \mathbb{A}} d_b \, \theta_{b,t} + e \, E(\theta).$$

For general economies, we provide some fundamental theoretical results: equilibria in pure strategies exist and they induce an ergodic stochastic process over the equilibrium configuration of actions. In finite economies, equilibria are unique and select a unique equilibrium in the infinite horizon limit. 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.[2] 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.

Importantly, we exploit our characterization results of the equilibria to address generally the issue of identification of social interactions in our context, with population data. In economies with social interactions, identification fundamentally entails distinguishing preferences for conformity (social interactions) from selection into social groups; see Manski (1993). Indeed, while a

[2]Standard variational arguments require bounding the marginal effect of any infinitesimal change dy_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 dy_a affects agent a's value function directly and indirectly, through its effects on all other agents's choices, which in turn affect agent a's value function. The marginal effect of any infinitesimal change dy_a is then an infinite sum of endogenous terms.

significant high correlation of socioeconomic choices across agents is often interpreted as evidence of social interactions, the spatial correlation of actions at equilibrium can be also due to the spatial correlation of preference shocks. More formally, take two agents, e.g., agent a and agent b. A positive correlation between $y_{a,t}$ and $y_{b,t}$ could be due to e.g., preferences for conformity. But the positive correlation between $y_{a,t}$ and $y_{b,t}$ could also be due to a positive correlation between $\theta_{a,t}$ and $\theta_{b,t}$. 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. High correlations of substance abuse between adolescent friends, for instance, could be due to social interactions or to friendship relations being selected in terms of demographic and psychological characteristics.

In the *dynamic linear conformity economies* studied in this paper, we show that dynamic forward looking behaviour plays an instrumental role in inducing identification.[3] We first show that exploiting the whole dynamic restrictions imposed by the model is not enough to obtain identification for general stochastic processes for preference shocks. We then show how specific relevant restrictions on the stochastic process, which are meant to capture natural properties of the selection mechanism inducing spatially correlated preferences, guarantee identification in our economy. More specifically, we prove that identification obtains if preference shocks are affected by a vector of observable time-varying individual characteristics, $x_{a,t}$ satisfying a *finite temporal memory* property:

$$\theta_{a,t} = \gamma x_{a,t} + \delta(x_{a-1,t} + x_{a+1,t}) + u_{a,t}, \quad E(x_{a,t}|x_{b,s}) = E(x_{a,t}), \text{ if } t > s + M, \ \forall a, b \in \mathbb{A};$$

where $u_{a,t}$ is an error term which is unobserved by the econometrician but observed by the agents. Intuitively, the fundamental effect of forward-looking behaviour in equilibrium consists in having current actions partly determined by expectations over future actions. Because of the general pattern on correlation over time and space of the stochastic process for the shocks, the future actions of any agent depend, in turn, on all the agents' future shocks. But under *finite temporal memory*, observables sufficiently in the past do not affect future shocks (and hence the expectations of future actions) and hence they can be used as valid instruments for identification.

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. Furthermore, our analysis extends to the addition of global interactions.[4] Finally, and perhaps most importantly, our analysis extends to encompass a richer structure of dynamic dependence of agents' actions in

[3] Interestingly, Manski (1993) alludes to the possibility that dynamic forward looking behaviour might help identification in linear models with social interactions.

[4] One particular form of global interactions occurs when each agent's preferences depend on an average of actions

equilibrium, e.g., economies in which agents' past behaviour is aggregated through an accumulated stock variable which carries habit persistence.[5]

1.1 Related literature

As we noted, the large majority of the existing models of social interactions are either static or myopic.[6] The general class of dynamic economies with social interactions we study in this paper are theoretically equivalent to a class of stochastic games, with an infinite number of agents, and uncountable state spaces. Available theoretical results for this class of games are extremely limited, even just regarding existence of "pure strategy" Markov-perfect equilibria; see Mertens and Parthasarathy (1987), Nowak (2003), and Duffie et al. (1994).[7] Bisin, Horst, and Özgür (2006) obtain existence and ergodicity results in this class of economies but only for the specific case of one-sided interactions across agents. This assumption greatly simplifies the analysis, as it excludes strategic interactions, but it is substantive as it severely limits the scope of social interactions to those which are structured hierarchically. Our work is related but complementary also to a body of work that studies dynamic discrete games of imperfect competition in empirical industrial organization[8] The restriction to linear economies we adopt in this paper has generally appealing analytical properties. Hansen and Sargent (2004) study this class of economies systematically, exploiting the tractability of linear control methods, but we are aware of only limited theoretical results regarding dynamic linear economies *with social interactions*: Bisin, Horst, and Özgür (2006) specialize their general economy to the linear case, but still only for the case of one-sided interactions; while Ioannides and Soetevent (2007) study dynamic linear economies with social interactions under the assumption that lagged rather than

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 behaviour, such as specific group cultures, or other externalities as well as price effects.

[5]These economies are naturally applied to the study of addiction. With respect to the addiction literature, as e.g., Becker and Murphy (1988), these economies treat peer effects not only in a single-person decision problem, but rather in a social equilibrium, allowing for the intertemporal feedback channel between agents across social space and through time; see also Becker et al. (1994), Gul and Pesendorfer (2007), Gruber and Koszegi (2001); see Rozen (2010) for theoretical foundations for intrinsic linear habit formation; see also Elster (1999) and Elster and Skog (1999) for surveys.

[6]Few exceptions include an example on female labor force participation in Glaeser and Scheinkman (2001), Binder and Pesaran (2001) on life-cycle consumption, Blume (2003) on social stigma, Brock and Durlauf (2010) and De Paula (2009) on duration models.

[7]See also Mertens (2002) and Vieille (2002) for surveys.

[8]See e.g. Aguirregabiria and Nevo (2013), Bajari, Hong, Nekipelov (2013) for good surveys. The typical model in this field has a finite number of agents, has a finite set of actions (see e.g. Bajari, Benkard, Levin (2007) for an exception), a finite-dimensional (typically finite) set of commonly observed states, and most importantly non-overlapping groups (e.g. geographically segmented markets).

contemporaneous average choices of peers affect individuals contemporaneous utility, thereby also greatly limiting the role of strategic interactions.

The analytical properties of linear economies which we exploit in our theoretical analysis come at a cost in terms of identification. In the context of linear economies, in fact, the issues pertaining to the distinction of preferences for conformity (social interactions) and correlated preferences across agents (selection into social groups) are fully general, running much deeper than Manski (1993)'s reflection problem.[9] In static linear economies, identification can be obtained if the population of agents can be collected into heterogeneous reference groups and under appropriate restrictions on the distribution of the agents' shocks; see Bramoullé et al. (2009) and Davezies et al. (2009) for identification in overlapping groups and Graham and Hahn (2005) for non-overlapping groups.[10] In dynamic economies with social interactions, identification is studied by De Paula (2009) for duration models. We are not aware of any results in the class of dynamic linear models with social interactions studied in this paper.

2 Dynamic Linear Conformity Economies

Time is discrete and is denoted by $t = 1, \ldots, T$. We allow both for infinite economies ($T = \infty$) and economies with an end period ($T < \infty$). A typical economy is populated by a set of *agents* $\mathbb{A}$, a generic element of which being denoted by a. We allow for both the case in which $\mathbb{A}$ is countably infinite and the case in which it is finite. In the latter case, it is convenient to dispose agents on a circle, to maintain symmetry; while in the countable limit, $\mathbb{A} := \mathbb{Z}$. Each agent lives for the duration of the economy.

At the beginning of each period t, agent a's random preference *type* $\theta_{a,t}$ is drawn from $\Theta = [\underline{y}, \bar{y}] \subset \mathbb{R}$, where $\underline{y} < \bar{y}$. Let $\theta := (\theta_t) := (\theta_{a,t})_{a \in \mathbb{A}, t \geq 1}$ be a stochastic process of agents' type profiles. After the realization of θ_t, agent $a \in \mathbb{A}$ chooses an *action* $y_{a,t}$ from the set $Y = [\underline{y}, \bar{y}]$. Let $\mathbf{Y} := \{y = (y_a)_{a \in \mathbb{A}} : y_a \in Y\}$ be the space of individual action profiles. Similarly, let $\mathbf{\Theta} := \{\theta = (\theta_a)_{a \in \mathbb{A}} : \theta_a \in \Theta\}$ be the space of individual type profiles. We assume, with no loss of generality, that the process $\theta = (\theta_1, \theta_2, \cdots)$ is defined, on the canonical probability space $(\Omega, \mathcal{F}, P)$, i.e., $\Omega := \{(\theta_1, \theta_2, \cdots) : \theta_t \in \mathbf{\Theta}\,, t = 1, 2, \cdots\}$. We equip the configuration spaces $\mathbf{Y}$ and $\mathbf{\Theta}$ with the product topologies, so that compactness of the individual action and type spaces implies compactness of the configuration spaces. The sequence $(\mathcal{F}_1, \mathcal{F}_2, \cdots)$ of Borel sub-σ-fields of $\mathcal{F}$ is a filtration in $(\Omega, \mathcal{F})$, that is $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}$. The process $\theta = (\theta_1, \theta_2, \cdots)$ is adapted to the

[9] See Brock and Durlauf (2001b) for conditions under which identification can instead be obtained in static non-linear models.

[10] Blume, Brock, Durlauf, and Ioannides (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. Other recent contributions include Blume, Brock, Durlauf, and Jayaraman (2015), Evans et al. (1992), Glaeser and Scheinkman (2001), Graham (2008), Ioannides and Zabel (2008), and Zanella (2007).

filtration $(\mathcal{F}_t : t \geq 1)$, i.e., for each t, θ_t is measurable with respect to $\mathcal{F}_t$. Finally, $P : \mathcal{F} \to [0,1]$ is a probability measure where $P((\theta_1, \dots, \theta_t) \in A) := P(\{\theta \in \Omega : (\theta_1, \dots, \theta_t) \in A\})$, all $A \in \mathcal{F}_t$. We require that $E[\theta_{a,s} \mid \theta_1, \dots, \theta_t] \in (\underline{y}, \bar{y})$, for any $t < s \leq T$ and any $a \in \mathbb{A}$, to guarantee that agents' choices be interior.

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 represents agent a's social reference group. The map $\mathbb{A} : N \to 2^{\mathbb{A}}$ 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, $\{y_{b,t}\}_{b \in N(a)}$, representing social interactions as direct preference externalities. Agent a's instantaneous preferences also depend on the agent's own previous choice, $y_{a,t-1}$, 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

$$\left(y_{a,t-1}, y_{a,t}, \{y_{b,t}\}_{b \in N(a)}, \theta_{a,t}\right) \mapsto u\left(y_{a,t-1}, y_{a,t}, \{y_{b,t}\}_{b \in N(a)}, \theta_{a,t}\right)$$

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

With the objective of providing a clean and simple analysis of dynamic social interactions in a conformity economy, we impose assumptions that are natural but stronger than required.[11] In particular (i) we restrict preferences to be quadratic, so as to restrict ourselves to a linear economy; (ii) we restrict the neighborhood correspondence to represent the minimal interaction structure allowing for overlapping groups; and we impose enough regularity conditions on the agents' choice problem to render it convex.[12] On the other hand, we do not impose any restriction on the correlation structure of the stochastic process θ.[13]

Assumption 1 *Let $\mathbb{A}$ represent the countable or finite (on a circle) set of agents:*

1. *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*

$$\begin{aligned} u(y_{a,t-1}, y_{a,t}, y_{a-1,t}, y_{a+1,t}, \theta_{a,t}) := & -\alpha_1 (y_{a,t-1} - y_{a,t})^2 - \alpha_2 (\theta_{a,t} - y_{a,t})^2 \\ & - \alpha_3 (y_{a-1,t} - y_{a,t})^2 - \alpha_3 (y_{a+1,t} - y_{a,t})^2 \end{aligned} \tag{1}$$

   where $\alpha_i \geq 0$, $i = 1, 2, 3$, and either α_1 or α_2 is strictly positive.

[11] See Section 2.3 for possible directions in which the structure and the results we obtain are easily generalized.

[12] While we model preferences for conformity directly as a preference externality, we intend this as a reduced form of models of behaviour in groups which induce indirect preferences for conformity, as e.g., Jones (1984), Cole et al. (1992), Bernheim (1994), Peski (2007).

[13] As we stated explicitly in the description of the environment on page 5, we only require that it satisfies $E[\theta_{a,s} \mid \theta_1, \dots, \theta_t] \in (\underline{y}, \bar{y})$, for any $t < s \leq T$ and any $a \in \mathbb{A}$, to guarantee that agents' choice be interior. In Section 3, we will show that identification does indeed require some assumption on the process θ.

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 $(y_{a,t-1}, \theta_{a,t})$ and the utility he receives from conforming to the current choices of his peers $(y_{a-1,t}, y_{a+1,t})$. Different values of α_i represent different levels of intensity of the social interaction motive relative to the own (or intrinsic) motive. The requirements that $\alpha_1 > 0$ or $\alpha_2 > 0$ anchor agents' preferences on their own private types or past choices. It is easy to see that, without such anchor, actions are driven only by social interactions and a large multiplicity of equilibria arises. Notice also that, under Assumption 1-2, utility functions with coefficients $\alpha_1, \alpha_2, \alpha_3$ and $\lambda\alpha_1, \lambda\alpha_2, \lambda\alpha_3$ represent the same preferences for any $\lambda > 0$. Hence restricting $\alpha_1+\alpha_2+2\alpha_3 = 1$ represents a normalization which we adopt without loss of generality.

Under Assumption 1, the agents' choice problem is convex. Let $y^{t-1} = (y_0, y_1, \ldots, y_{t-1})$ and $\theta^{t-1} = (\theta_1, \ldots, \theta_{t-1})$ be the choices and type realizations up to period $t-1$, where $y_0 \in \mathbf{Y}$ is the initial configuration. Before each agent's time t choice, both y^{t-1} and θ^t are observed by all agents.

A *strategy* for an agent a is a sequence of functions $y_a = (y_{a,t})$, adapted to the filtration $\mathcal{F}$, where for each t, $y_{a,t} : \mathbf{Y}^t \times \mathbf{\Theta}^t \to Y$. 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 y_a to maximize

$$E\left[\sum_{t=1}^{T} \beta^{t-1} u\left(y_{a,t-1}, y_{a,t}, \{y_{b,t}\}_{b\in N(a)}, \theta_{a,t}\right) \,\Big|\, (y_0, \theta_1)\right] \tag{2}$$

given the strategies of other agents and given $(y_0, \theta_1) \in \mathbf{Y} \times \mathbf{\Theta}$.

Note that Assumption 1-1 imposes enough symmetry on the agents' problem so that it is enough to analyze the optimization problem relative to a single reference agent, say agent $0 \in \mathbb{A}$. Indeed, 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 $b-a$.[14] In the finite-horizon economy, at each time t, let $l = T-(t-1)$ denote the time periods remaining until the end of the economy T. The optimal choice of any economic agent $a \in \mathbb{A}$ at t, for any $t \in \{1, \ldots, T\}$, is then determined by a continuous choice function $g_l : \mathbf{Y} \times \mathbf{\Theta}^t \to Y$:

$$y_{a,t}\left(g\right)\left(y^{t-1}, \theta^t\right) = g_l(R^a\, y_{t-1}, R^a\, \theta^t)$$

In the infinite-horizon economy, the optimal choice of any economic agent $a \in \mathbb{A}$ at any t, is instead stationary and determined by a continuous choice function $g : \mathbf{Y} \times \mathbf{\Theta}^t \to Y$, that is:

$$y_{a,t}\left(g\right)\left(y^{t-1}, \theta^t\right) = g(R^a\, y_{t-1}, R^a\, \theta^t)$$

[14] That is, $c \in N(a)$ if and only if $c + (b-a) \in N(b)$. The operations of addition and subtraction are legitimate for $\mathbb{A} := \mathbb{Z}$ as well as for $\mathbb{A}$ represented by a circle, defined modularly.

Definition 1 *A symmetric subgame perfect equilibrium of a* dynamic linear finite-horizon conformity economy *is a sequence of maps* $(g_l^*)_{l=T}^1 : \mathbf{Y} \times \mathbf{\Theta}^t \to Y$ *adapted to* $\mathcal{F}$ *such that for all* $a \in \mathbb{A}$ *and for all* $(y^{t-1}, \theta^t) \in \mathbf{Y}^t \times \mathbf{\Theta}^t$

$$g_l^*\left(R^a\, y_{t-1}, R^a\, \theta^t\right) \in \underset{y_{a,t} \in Y}{\operatorname{argmax}}\; E\left[\sum_{t=1}^{T} \beta^{t-1} u\left(y_{a,t-1}, y_{a,t}, \{y_{b,t}(g^*)\}_{b \in N(a)}, \theta_t^a\right) \Big| (y_0, \theta_1)\right].$$

A symmetric subgame perfect equilibrium of a dynamic linear infinite-horizon conformity economy *is a stationary map* $g^* : \mathbf{Y} \times \mathbf{\Theta}^t \to Y$ *adapted to* $\mathcal{F}$ *such that for all* $a \in \mathbb{A}$ *and for all* $(y^{t-1}, \theta^t) \in \mathbf{Y}^t \times \mathbf{\Theta}^t$

$$g^*\left(R^a\, y_{t-1}, R^a\, \theta^t\right) \in \underset{y_{a,t} \in Y}{\operatorname{argmax}}\; E\left[\sum_{t=1}^{T} \beta^{t-1} u\left(y_{a,t-1}, y_{a,t}, \{y_{b,t}(g^*)\}_{b \in N(a)}, \theta_t^a\right) \Big| (y_0, \theta_1)\right].$$

2.1 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. Unless otherwise mentioned specifically, the proofs of all statements and other results can be found in the Supplemental Appendix.

Theorem 1 (Existence) *Consider a* dynamic linear conformity economy*.*

(i) If the time horizon is finite ($T < \infty$)*, the economy admits a unique subgame perfect equilibrium which is symmetric, with*

$$\begin{aligned} g_l^*(R^a\, y_{t-1},\ R^a \theta^t) \;\; &= \;\; \sum_{b \in \mathbb{A}} c_{b,l}\, y_{a+b,t-1} + \sum_{b \in \mathbb{A}} d_{b,l}\, \theta_{a+b,t} \\ &\quad + \sum_{b \in \mathbb{A}} \sum_{\tau=t+1}^{T} e_{b,l,\tau-t}\, E\left[\theta_{a+b,\tau} | \theta^t\right] \qquad P - a.s. \end{aligned}$$

Furthermore, $c_{b,l}, d_{b,l}, e_{b,l,\tau-t} \geq 0$, $\forall \tau = t+1, \ldots, T$, *and* $\sum_{b \in \mathbb{A}} (c_{b,l} + d_{b,l} + e_{b,l}) = 1$, *where* $e_{b,l} = \sum_{\tau=t+1}^{T} e_{b,l,\tau-t}$.

(ii) If the time horizon is infinite ($T = \infty$)*, the economy admits a symmetric stationary subgame perfect equilibrium, with*

$$g^*(R^a\, y_{t-1},\ R^a\, \theta^t) = \sum_{b \in \mathbb{A}} c_b\, y_{a+b,t-1} + \sum_{b \in \mathbb{A}} d_b\, \theta_{a+b,t} + \sum_{b \in \mathbb{A}} \sum_{\tau=t+1}^{\infty} e_{b,\tau-t}\, E\left[\theta_{a+b,\tau} | \theta^t\right] \qquad P - a.s.$$

Furthermore, $c_b, d_b,\ e_{b,\tau-t} \geq 0,\ \forall\tau = t+1,\ldots,\infty,$ *and* $\sum_{b\in\mathbb{A}}(c_b + d_b + e_b) = 1,$ *where* $e_b = \sum_{\tau=t+1}^{\infty} e_{b,\tau-t}$.[15]

(iii) g_l^* *converges to* g^* *as* $l \to \infty$*, in the sense that* $\lim_{l\to\infty}(c_{b,l}, d_{b,l}, e_{b,l}) = (c_b, d_b, e_b),\ \forall b \in \mathbb{A}$.

The proof of the existence theorem requires some subtle arguments. While referring to Appendix A 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 dy_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 dy_a affects agent a's value function not only directly, but also indirectly, that is through its effects on all agents $b \in \mathbb{A}\backslash\{a\}$'s choices, which in turn affect agent a's value function. The marginal effect of any infinitesimal change dy_a is then an infinite sum, and each term of the sum consists in turn of an infinite sum of endogenous marginal effects from all agents $b \in \mathbb{A}\backslash\{a\}$'s policy functions.[16] In our economy, with quadratic utility, policy functions g_l^* are obviously necessarily linear. 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.

Remark (Uniqueness). Social interaction economies are usually plagued with multiple equilibria. Previous uniqueness results in the literature require some form of '*Moderate Social Influence*' assumption, which roughly means that the effect on marginal utility of a change in individual's own choice should dominate the sum of the effects on marginal utility of changes in peers' choices (see e.g. Glaeser and Scheinkman (2003)). In fact, as long as either one of the parameters α_1 or α_2 is positive (which is true by Assumption 1), this is the case for the economies we study. More specifically, consider first a finite-horizon economy. The coefficients $c_{b,l}$ and $d_{b,l}$ satisfy

$$\lim_{|b|\to\infty} c_{a+b,l} = \lim_{|b|\to\infty} d_{a+b,l} = 0$$

The impact of an agent $a+b$ on agent a tends to zero as $|b| \to \infty$. In this sense, linear conformity economies display weak social interactions. This is why, in contrast to other models in the literature, no matter how large the interaction parameter α_3 is relative to the others, the equilibrium stays unique for finite-horizon economies. Furthermore, in the infinite horizon economy, while equilibria are not necessarily unique, there is a unique equilibrium which is the limit of equilibria in finite economies.

[15]Several assumptions can be relaxed while guaranteeing existence. This is the case, in particular, for the symmetry of the neighborhood structure; see Section 2.3.1 for the discussion.

[16]The methodology used by Santos (1991) to prove the smoothness of the policy function in infinite dimensional recursive choice problems also does not apply.

By exploiting the linearity of the 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. Let $C_l := \sum_b c_{b,l}$ denote the total effect of past actions; that is, the effect on an agent a's action of a uniform unitary increase in all agents' past actions. Similarly, let $D_l := \sum_b d_{b,l}$ denote the total effect of contemporary preference shocks; and $E_l := \sum_b e_{b,l}$ the total effect of expected future preference shocks.

Remark (Myopic economy) It is of interest to compare subgame perfect equilibria of dynamic linear conformity economies with equilibria of myopic (static) conformity economies. Myopic economies have in fact been extensively studied in the theoretical and empirical literature, following the mathematical physics literature in statistical mechanics on interacting particle systems.[17] More specifically, myopic behaviour in the literature is commonly characterized by the assumption that an agent $a \in \mathbb{A}$, when choosing $y_{a,t}$ at time t, does not consider choosing again in the future, nor neighbors choosing again in the future.[18] It follows straightforwardly that the coefficients of the policy function in these economies are equal to the ones of the unique subgame perfect equilibrium policy function of a dynamic linear conformity economy with finite horizon T with $\beta = 0$ (or, also, in the last period, at $t = T$). Therefore, a *linear conformity economy with myopic agents* admits a unique equilibrium which is symmetric with

$$g^*(R^a\, y_{t-1},\ R^a\, \theta^t) = \sum_{b \in \mathbb{A}} c_b\, y_{a+b,t-1} + \sum_{b \in \mathbb{A}} d_b\, \theta_{a+b,t} \qquad P - a.s.$$

In Sections 3.1 and 3.2, we will study whether the behaviour in the dynamic linear conformity economies, with rational forward looking agents, can be identified from behaviour in the same economies with myopic agents.

Theorem 2 (Characterization and computation) *Consider a T-period* dynamic linear conformity economy*.*

(i) The equilibrium coefficients, $(c_{b,l}, d_{b,l}, e_{b,l})$, for any $b \in \mathbb{A}$, $T \geq l \geq 1$, are independent of the properties of the stochastic process θ.

(ii) In the terminal time period, $l = 1$, the sequence of equilibrium coefficients $(c_{b,1}, d_{b,1}, e_{b,1})$, is

[17] See e.g., Blume and Durlauf (2001), Brock and Durlauf (2001b), Glaeser and Scheinkman (2003), and Özgür (2010) for a comprehensive survey; and Liggett (1985) for the mathematical literature.

[18] In some of the literature, myopic behaviour is modelled not only by assuming that all agents in the economy only interact once, but also that their neighbors do not not change their previous period actions. In this case, the dynamics describe a backward looking behaviour of the agents and it can be shown that the ergodic stationary distribution of actions coincides with that of *myopic* agents as characterized in the text; see e.g., Glaeser and Scheinkman (2003).

exponential in $|b|$.[19]

(iii) *For any* $\beta > 0$, *and any* $l = 2, \dots, \infty$, *the sequence of equilibrium coefficients* $(c_{b,l}, d_{b,l}, e_{b,l})$, *declines faster than exponentially in* $|b|$.[20] *It declines exponentially if* $\beta = 0$.

(iv) *Total effects of past actions and of contemporary preference shocks,* C_l *and* D_l *respectively, increase as* l *decreases; the total effect of expected future preference shocks,* E_l *instead decreases as* l *decreases.*[21]

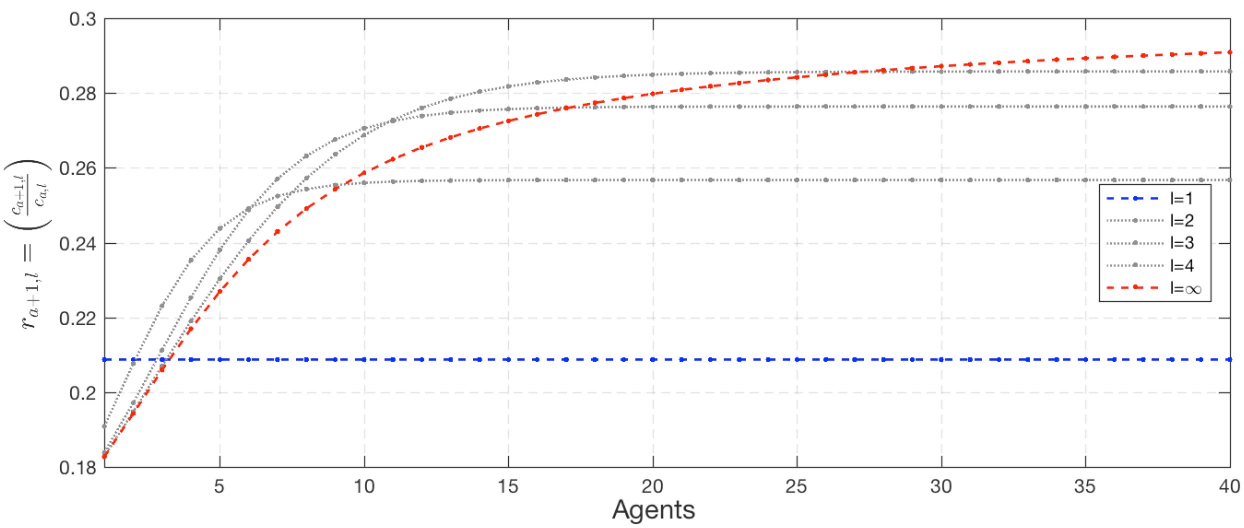

Figure 1: Cross-sectional rates of convergence.[22]

[19]Specifically, for any $b \in \mathbb{A}$,

$$c_{b,1} = r_1^{|b|} \left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right) \left(\frac{1 - r_1}{1 + r_1}\right) \quad \text{and} \quad d_{b,1} = r_1^{|b|} \left(\frac{\alpha_2}{\alpha_1 + \alpha_2}\right) \left(\frac{1 - r_1}{1 + r_1}\right)$$

where $r_1 = \left(\frac{\Delta_1}{2\,\alpha_3}\right) - \sqrt{\left(\frac{\Delta_1}{2\,\alpha_3}\right)^2 - 1}$ and $\Delta_1 = \alpha_1 + \alpha_2 + 2\,\alpha_3$. Note that $e_{b,1} = 0$ for all $b \in \mathbb{A}$ optimally since there are no future shocks.

[20]The sequence of equilibrium coefficients $(c_{b,l}, d_{b,l}, e_{b,l})$, are computed recursively iterating a map $L_l : \Delta_{c,d,e} \to \Delta_{c,d,e}$, $l = 2, \dots, T$, obtained from agent 0's dynamic program's first-order condition and characterized in (A.5), and (A.6).

[21]Specifically, C_l, D_l, and E_l take the form of continued fractions:

$$\begin{aligned} C_l &= \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1 \beta (1 - C_{l-1})} \\ D_l &= \frac{\alpha_2}{\alpha_1 + \alpha_2 + \alpha_1 \beta (1 - D_{l-1})} \\ E_l &= \frac{\alpha_1 \beta (1 - D_{l-1})}{\alpha_1 + \alpha_2 + \alpha_1 \beta (1 - D_{l-1})} \end{aligned}$$

with $C_1 = \frac{\alpha_1}{\alpha_1 + \alpha_2}$; $D_1 = \frac{\alpha_2}{\alpha_1 + \alpha_2}$; $E_1 = 0$.

[22]Parameter values used for this figure are $(\alpha_1, \alpha_2, \alpha_3, \beta) = (0.2, 0.2, 0.3, 0.95)$.

Result (i), that equilibrium coefficients are independent of the properties of θ, is a direct consequence of the linearity of policy functions. It allows us to study theoretically economies which are fully general in terms of the stochastic process for the shocks. Restrictions will however be required to guarantee identification, in Section 3. Results (ii) and (iii) imply that the sequence of equilibrium coefficients $(c_{b,l}, d_{b,l}, e_{b,l})$, converges in the cross-sectional tail at a constant rate for $l = 1$ and at an increasing rate for $\beta > 0$, $l \geq 2$; see Figure 1. This is a fundamental equilibrium implication of a dynamic economy (in a terminal period, in fact, the economy would be effectively static). It is instrumental in allowing us to identify forward looking from myopic behaviour in a dynamic economy; see Remark (Myopic behaviour) in Section 3. Finally, result (iv) on the dependence of total effects on the length of the remaining economy l, implies that in earlier periods agents put relatively more weight on expectations and, as the horizon gets shorter, they eventually shift weight on to history and current shocks. Importantly, while we do not have a formal result, simulations show that $(\frac{c_{0l}}{C_l})$, the relative weight on own history, decreases as l decreases.[23]

We turn now to the relationship between the equilibrium coefficients and the preference parameters of the economy. Fix the normalization $\alpha_1 + \alpha_2 + 2\alpha_3 = 1$, and consider α_2 determined residually by the normalization.

Theorem 3 (Dependence on the preference parameters) *Consider a* dynamic linear conformity economy *with $T > 1$ and $\alpha_1 \neq 0$.*

(i) At a Subgame Perfect equilibrium of a finite horizon economy, the map between parameters $(\alpha_1, \alpha_3, \beta)$ and coefficients $(c_{b,l})_{l=1,\, b\in\mathbb{A}}^{T}$ is injective. Similarly, at a subgame perfect equilibrium of an infinite horizon economy, the map between parameters $(\alpha_1, \alpha_3, \beta)$ and coefficients $(c_b)_{b\in\mathbb{A}}$ is injective.

(ii) Seeing the sequences of equilibrium coefficients $c_{b,l}$, $d_{b,l}$, and $e_{b,l}$ as distributions over b, an increase in α_3, keeping β and the ratio $\frac{\alpha_1}{\alpha_2}$ intact, induces a mean preserving spread in each of these distributions.[24]

Result (i) is a fundamental component of our identification results in Section 3. It reduces the problem of identifying $(\alpha_1, \alpha_3, \beta)$ to one of consistently estimating the parameters $(c_{b,l})_{l=1,\, b\in\mathbb{A}}^{T}$ (or $(c_b)_{b\in\mathbb{A}}$, when the horizon is infinite). Result (ii) is represented in Figure 2. It implies that, by weighting more the action of closer neighbors, a large α_3 induces equilibrium actions which are more concentrated along the social space. This in turn can be seen in Figure 3.[26]

[23]The structural dependence of the parameters' configuration on the length of the remaining economy l is in principle useful to distinguish the implications of social interactions from those e.g., of habit persistence models, that also generally induce interesting non-stationary effects.

[24]Indeed, total effects C_l, D_l, E_l are independent of α_3.

[26]See Appendix J for details of the simulations generating the figures.

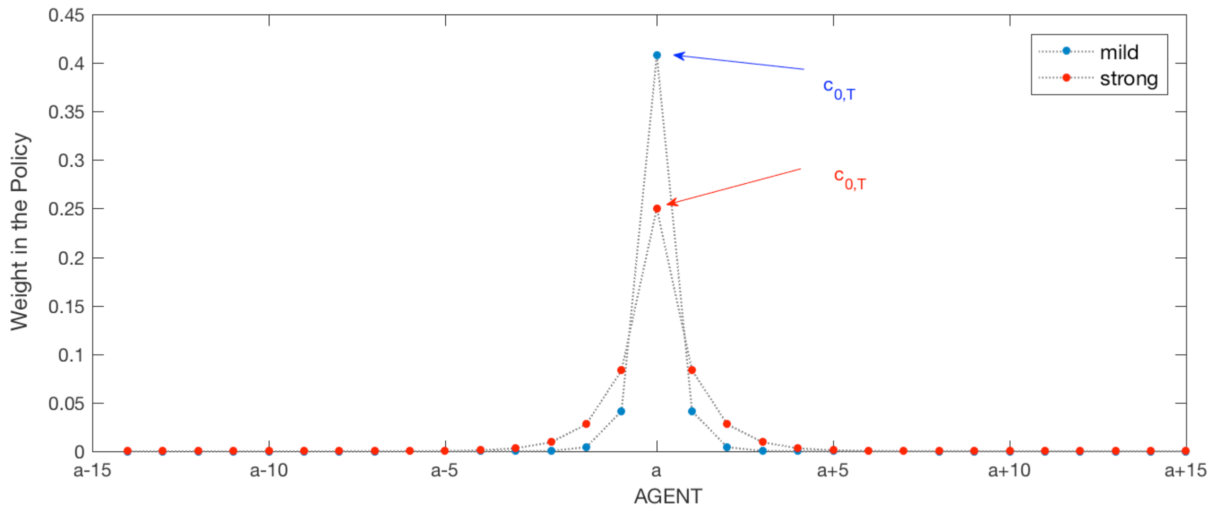

Figure 2: Distribution of policy coefficients on past actions for mild and strong interactions.[25]

Remark (Markov economy) A *(s-)Markov linear conformity economy*, for $s \in \mathbb{N}$, is a *dynamic linear conformity economy* whose stochastic process θ is (s-)Markov, that is, it satisfies $P(\{\theta \in \Omega : \theta^t \in A\}) = P(\{\theta' \in \Omega : \theta'^t \in A\})$ for all θ, θ' such that $(\theta_{t-s}, \theta_{t-s-s}, \ldots, \theta_t) = (\theta'_{t-s}, \theta'_{t-s-s}, \ldots, \theta'_t)$, and for all $A \in \mathcal{F}_t$.

If the time horizon is finite ($T < \infty$), the unique subgame perfect equilibrium in a *(s-)Markov linear conformity economy* is (s-)Markov, with

$$\begin{aligned} y_{a,t} \; = \; g_l^*(R^a\, y_{t-1},\; R^a\,(\theta_{t-s}, \ldots, \theta_t)) \;\; &= \;\; \sum_{b\in\mathbb{A}} c_{b,l}\, y_{a+b,t-1} + \sum_{b\in\mathbb{A}} d_{b,l}\, \theta_{a+b,t} \\ &+ \sum_{\tau=t+1}^{T} \sum_{b\in\mathbb{A}} e_{b,l,\tau-t}\, E\left[\theta_{a+b,\tau} | \theta_{t-s}, \ldots, \theta_t\right] \qquad P-a.s. \end{aligned} \tag{3}$$

If the time horizon is infinite ($T = \infty$), there exists a (s-)Markov symmetric subgame perfect equilibrium, with

$$\begin{aligned} y_{a,t} \; = \; g^*(R^a\, y_{t-1},\; R^a\; (\theta_{t-s}, \ldots, \theta_t)) \;\; &= \;\; \sum_{b\in\mathbb{A}} c_b\, y_{a+b,t-1} + \sum_{b\in\mathbb{A}} d_b\, \theta_{a+b,t} \\ &+ \sum_{\tau=t+1}^{\infty} \sum_{b\in\mathbb{A}} e_{b,\tau-t}\, E\left[\theta_{a+b,\tau} | \theta_{t-s}, \ldots, \theta_t\right] \qquad P-a.s. \end{aligned} \tag{4}$$

such that $c_{b,l}, d_{b,l}, e_{b,l}$ converge to c_b, d_b, e_b. Once again, in the infinite horizon economy, when equilibria are not necessarily unique, we can select the unique limit of finite economies.

[26] Final period policy is used. Parameter values used are $\alpha_3 = 0.1$ for *Mild* and $\alpha_3 = 0.3$ for *Strong* interactions. The rest of the parameters are $\alpha_1 = \alpha_2 = 0.5(1 - 2\alpha_3)$ and $\beta = 0.95$. See Appendix J for details of the simulations.

Remark (Social welfare)[27] Social interactions are modeled in this paper as a preference externality, that is, by introducing a dependence of agent a's preferences on his/her peers' actions. A benevolent social planner, taking into account the preference externalities and at the same time treating all agents symmetrically, maximizes the expected discounted utility of a generic agent, say of agent $a \in \mathbb{A}$, by choosing a symmetric choice function h in the space of bounded, continuous, and Y-valued measurable functions $CB(\mathbf{Y} \times \mathbf{\Theta}, Y)$. The choice of h induces, in a recursive way, a sequence of choices for any agent $b \in \mathbb{A}$, given (y_0, θ_1). A subgame perfect equilibrium of a conformity economy is inefficient, in the sense that the sequence of choices induced by the benevolent planner for each agent Pareto-dominate the equilibrium choices. Most importantly, an efficient policy function will weight less heavily the agent's own-effect and more heavily other agents' effects, relative to the equilibrium policy.

2.2 Ergodicity

Given the characterization of the parameters of the policy function at hand, we are also able to determine the long-run behaviour of the equilibrium process emerging from the class of dynamic models we study. To that end, let an infinite-horizon economy with conformity preferences be given and let g^* be a symmetric subgame perfect equilibrium (recall that Theorem 1 does not guarantee that a unique such g^* exists). Let π_0 be an initial distribution on the space of action profiles $\mathbf{Y}$. Given π_0 and a stationary (s-)Markov process for (θ_t), g^* induces an equilibrium process $(y_t \in Y)_{t=0}^{\infty}$ and an associated transition function Q_{g^*}. This latter generates iteratively a sequence of distributions $(\pi_t)_{t=1}^{\infty}$ on the configuration space $\mathbf{Y}$, i.e., for $t = 0, 1, \ldots$

$$\pi_{t+1}(A) = \pi_t Q_{g^*}(A) = \int_{\mathbf{Y}} Q_{g^*}(y_t \ldots, y_{t-s}, A)\, \pi_t(dy_{t+1})$$

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 equilibrium process starting from π is *ergodic*.[28]

Furthermore, we show that, for any initial distribution π_0 and a (s-)Markov symmetric subgame perfect equilibrium policy function g^*, the equilibrium process $(y_t \in Y)_{t=0}^{\infty}$ converges in distribution to the invariant distribution π, independently of π_0. This also implies that π is the *unique* invariant distribution of the equilibrium process $(y_t \in Y)_{t=0}^{\infty}$. More specifically,

Theorem 4 (Ergodicity) *Consider a* dynamic linear conformity economy*. The equilibrium process $(y_t \in Y)_{t=0}^{\infty}$ induced by a stationary (s-)Markov symmetric subgame perfect equilibrium*

[27] We relegate the formal treatment of Social welfare to the Technical Appendix H.

[28] We call a Markov process (y_t) with state space $\mathbf{Y}$ under a probability measure P ergodic if $\frac{1}{T}\sum_{t=1}^{T} f(y_t) \to \int f dP$ P-almost surely for every bounded measurable function $f : \mathbf{Y} \to \mathbb{R}$. See, e.g., Duffie et al. (1994) for a similar usage.

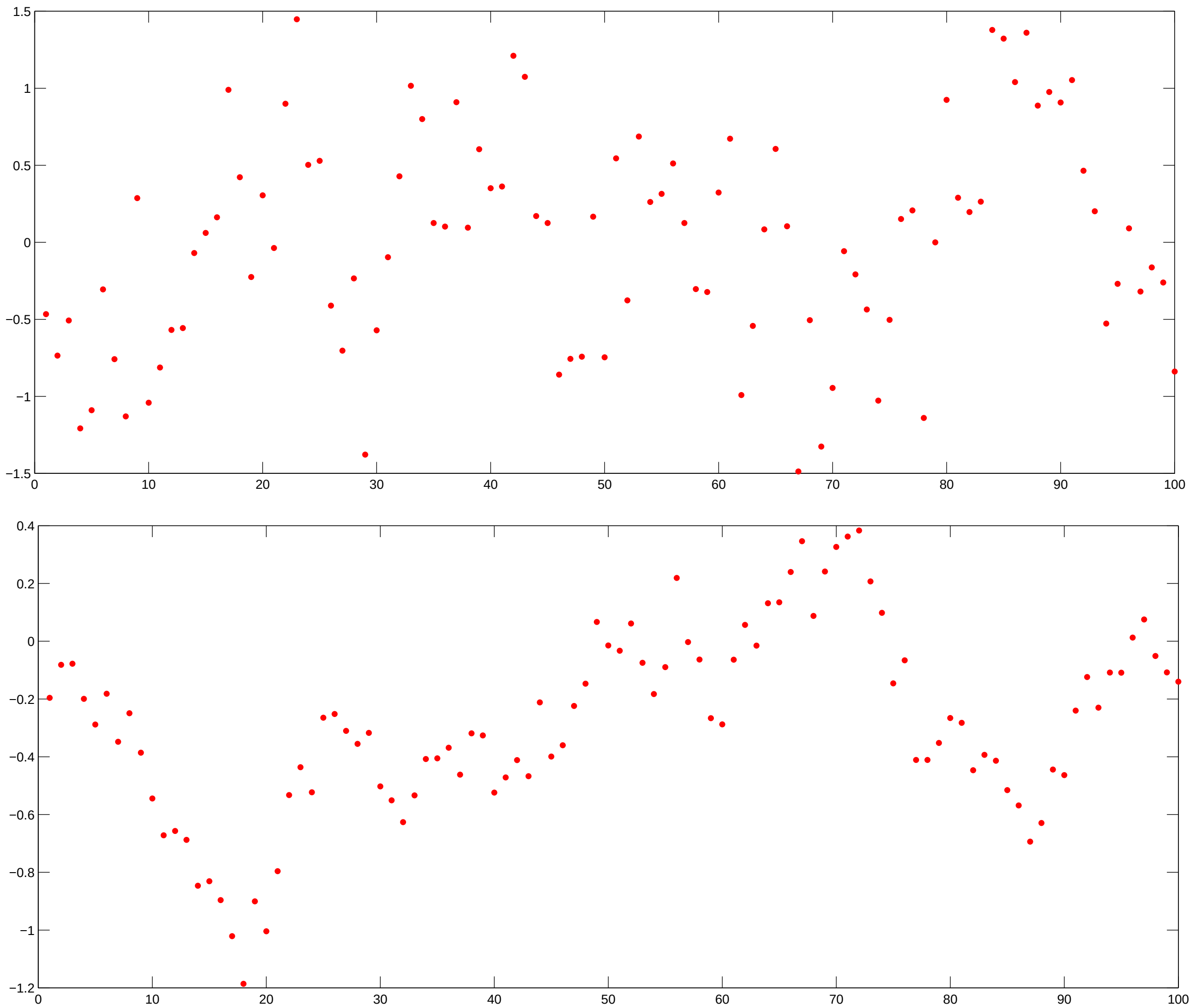

Figure 3: Simulated stationary profile of actions in a dynamic economy with *Mild* (top) and *Strong* (bottom) α_3. The horizontal axis represents agents and the vertical axis represents agents' choices from a common interval. Infinite-horizon policy is iterated with i.i.d. shocks and a large number of agents. Parameter values used are $\alpha_3 = 0.1$ for *Mild* and $\alpha_3 = 0.3$ for *Strong* interactions. The rest of the parameters are $\alpha_1 = \alpha_2 = 0.5(1 - 2\alpha_3)$ and $\beta = 0.95$. See Appendix J for details of the simulations.

via the policy function $g^*(y_{t-1},\ (\theta_{t-s},\dots,\theta_t))$ and the unique invariant measure π as the initial distribution is ergodic; π is the joint distribution of

$$y_t = \Bigg(\sum_{m=1}^{\infty}\sum_{b_1}\cdots\sum_{b_m} c_{b_1}\cdots c_{b_{m-1}}\Big[d_{b_m}\,\theta_{a+b_1+\cdots+b_m,t+1-m} + \sum_{\tau=1}^{\infty} e_{b_m,\tau}\,E[\theta_{a+b_1+\ldots+b_m,t+1-m+\tau}\,|\,\theta_{1-m},\dots,\theta_{1-m-s}]\Big]\Bigg)_{a\in\mathbb{A}} \tag{5}$$

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 $(y_t \in Y)_{t=0}^{\infty}$ converges to π in the topology of weak convergence for probability measures, independently of any initial distribution π_0.[29]

2.3 Extensions

The class of social interaction economies we study is restricted along several dimensions. Some of these restrictions, however, might turn out to be important in applications and empirical work. In this section we briefly 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. For simplicity we consider here only $(s-)$Markov linear conformity economies with infinite time horizon, $T=\infty$. Extension to general stochastic processes for θ and/or to finite time horizons are straightforward.[30]

2.3.1 General Network Structures

Consider a linear conformity economy with arbitrary neighborhood structure (not necessarily translation invariant), $N : \mathbb{A} \to 2^{\mathbb{A}}$. In particular, let a generic agent a's preferences for conformity be represented by a general term $-\sum_{b\in N(a)} \alpha_{3,a,b}\,(y_{b,t} - y_{a,t})^2$, entering additively in the utility function u_a.

At a $(s-)$Markov Subgame Perfect equilibrium of this economy, policy functions satisfy

$$\begin{aligned} y_{a,t} \;&=\; \sum_{b\in\mathbb{A}} c_{a,b}\,y_{b,t-1} + \sum_{b\in\mathbb{A}} d_{a,b}\,\theta_{b,t} \\ &\quad + \sum_{\tau=t+1}^{\infty}\sum_{b\in\mathbb{A}} e_{a,b,\tau-t}\,E\left[\theta_{b,\tau}|\theta_{t-s},\dots,\theta_t\right] \qquad P-a.s.; \end{aligned} \tag{6}$$

[29] A 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 $f:\mathbf{Y}\to\mathbb{R}$, $\lim_{t\to\infty}\int f\,d\lambda_t = \int f d\lambda$ almost surely (see e.g. Kallenberg (2002), p.65).

[30] We sketch proof arguments for all extensions in Section I of the Technical Appendix.

and our analysis extends.[31]

2.3.2 Social Accumulation of Habits

Consider a linear conformity economy where individual behaviour depends on the accumulated stock of present and previous choices

$$r_{a,t} = (1-\delta)\, r_{a,t-1} + y_{a,t}$$

rather than on her present choice $y_{a,t}$ only. In particular, let a generic agent a's preferences for conformity be represented by a general term $-\alpha_1 (r_{a,t} - y_{a,t})^2$, entering additively in the utility function u_a.[32]

At a ($s-$)Markov subgame perfect equilibrium of this economy, policy functions satisfy

$$\begin{aligned} y_{a,t} &= \sum_{b\in\mathbb{A}} c_b\, r_{a+b,t-1} + \sum_{b\in\mathbb{A}} d_b\, \theta_{a+b,t} \\ &+ \sum_{\tau=t+1}^{\infty}\sum_{b\in\mathbb{A}} e_{b,\tau-t}\, E\left[\theta_{a+b,\tau}|\theta_{t-s},\ldots,\theta_t\right] \qquad P-a.s.; \end{aligned} \tag{7}$$

and our analysis extends.

2.3.3 Global Interactions

Consider a linear conformity economy where individual behaviour is affected by global as well as local determinants.[33] 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 y be defined as

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

[31] However, in this economy, for finite time horizon $T < \infty$, uniqueness of subgame perfect equilibrium is not guaranteed with an infinite number of agents. It is however sufficient for uniqueness that the relative composition of the peer effects within the determinants of individual choice be uniformly bounded:

$$\exists\ 0 < K < 1 \text{ such that}, \forall a \in \mathbb{A},\ \ \frac{\sum_{b\in N(a)} \alpha_{3,a,b}}{\alpha_1 + \alpha_2 + \sum_{b\in N(a)} \alpha_{3,a,b}} < K;$$

see e.g. Glaeser and Scheinkman (2003), Horst and Scheinkman (2006), and Ballester, Calvó-Armengol, and Zenou (2006) for *Moderate Social Influence* conditions restrictions in a similar spirit.

[32] For instance, $r_{a,t}$ captures what the addiction literature calls a "*reinforcement effect*" on agent a's substance consumption.

[33] With respect to 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 a generic agent a's preferences for conformity be represented by a general term $-\alpha_4 \left(p(y_t) - y_{a,t}\right)^2$, entering additively in the utility function u_a.

At a ($s-$)Markov subgame perfect equilibrium of this economy, policy functions satisfy

$$y_{a,t} = \sum_{b \in \mathbb{A}} c_b \, y_{a+b,t-1} + \sum_{b \in \mathbb{A}} d_b \, \theta_{a+b,t} + \sum_{\tau=t+1}^{\infty} \sum_{b \in \mathbb{A}} e_{b,\tau-t} \, E\left[\theta_{a+b,\tau} | \theta_{t-s}, \ldots, \theta_t\right] + B(p(y_1) \qquad P - a.s., \tag{8}$$

for some constant $B(p(y_1))$ that depends only on the initial average action, $p(y_1)$ (which is assumed to exist); and our analysis extends.

3 Identification

We study here the identification properties of *dynamic linear conformity economies.* Identification obtains when the restrictions imposed on equilibrium choices allow to unambiguously recover the model's parameters from observed actions. In economies with social interactions, this fundamentally entails distinguishing preferences for conformity (social interactions) and correlated preferences across agents, which can be generally due to some sort of assortative matching or positive selection into social groups. A positive correlation between $y_{a,t}$ and $y_{b,t}$, for any two agents a and b at some time t, could be due to social interaction preferences, with $\alpha_3 > 0$, or to a positive correlation between $\theta_{a,t}$ and $\theta_{b,t}$ even in the presence of no social interactions, with $\alpha_3 = 0$.

Formally, we assume that the time horizon $T \geq 2$[34] and the size of the economy N are known (to the econometrician as well as to the agents); and so is the population probability distribution over actions $F(y)$, e.g., because an arbitrarily large number of replications of the economy is observed. Preference shocks may depend on observable covariates and, in that case, the probability distribution over these covariates is also known. The structural parameters of the model to be recovered are: the weights placed on own past action α_1, own preference shock α_2, and others' current actions in current utility α_3, as well as the discount factor β. Fixing the normalization $\alpha_1 + \alpha_2 + 2\alpha_3 = 1$, we focus on α_1, α_3, and β as the parameters of interest. Denote by $F_{\alpha_1,\alpha_3,\beta}(y)$ the population probability distribution under parameters α_1, α_3 and β. Identification means that these structural parameters can be uniquely recovered from the probability distribution on outcomes.[35]

Definition 2 *A* dynamic linear conformity economy *is identified if $F_{\alpha_1,\alpha_3,\beta} = F_{\alpha_1',\alpha_3',\beta'}$ implies that $\alpha_1 = \alpha_1'$, $\alpha_3 = \alpha_3'$ and $\beta = \beta'$.*

[34] The static case $T = 1$ is degenerate in our context.

[35] The notion of identification we adopt is parametric. More generally, Rust (1994) and Magnac and Thesmar (2002) show that, even in dynamic discrete choice models, utility functions cannot be non-parametrically identified.

The properties of the stochastic process of preference shocks $(\theta_{a,t})_{a\in\mathbb{A},t\geq 1}$ clearly affect identification. We prove, first, that the model is not identified when no restrictions are placed on this process. We then prove our main result in this section, determining conditions on the structure of preference shocks, and on how they depend on observables, under which identification holds.

Theorem 5 *A* dynamic linear conformity economy *with no restriction on the stochastic process of preference shocks* $(\theta_{a,t})_{a\in\mathbb{A},t\geq 1}$ *is not identified, for finite or infinite* T *and finite or infinite* N.

Proof: Suppose first that $T=\infty$. Consider the stationary distribution of a linear conformity economy with social interactions, that is, $\alpha_3>0$, and an i.i.d. preference shock process $(\theta_{a,t})_{a\in\mathbb{A},t\geq 1}$, where $\bar{\theta}$ is their identical expected value. We have shown in Section 2.2 that such stationary distribution is given by the ergodic measure π in (5), which under the i.i.d assumption takes the form

$$y_t=\left(\frac{E(\alpha)\,\bar{\theta}}{1-C(\alpha)}+\sum_{m=1}^{\infty}\sum_{b_1}\cdots\sum_{b_m}c(\alpha)_{b_1}\cdots c(\alpha)_{b_{m-1}}\left(d(\alpha)_{b_m}\;\theta_{a+b_1+\cdots+b_m,t+1-m}\right)\right)_{a\in\mathbb{A}}.$$

Consider now an alternative specification of this economy with *no interactions* between agents $(\alpha_3'=0)$ and *no habits* $(\alpha_1'=0)$ but simply a preference shock process $\{\theta_{a,t}'\}_{a\in\mathbb{A},t\geq 1}$ and *own type effects* with $\alpha_2'>0$. For this economy, equilibrium choice of agent a at time t is given by $y_{a,t}=\theta_{a,t}'$. As long as the process $\{\theta_{a,t}'\}_{a\in\mathbb{A},t\geq 1}$ is the one where

$$\theta_{a,t}':=\frac{E(\alpha)\,\bar{\theta}}{1-C(\alpha)}+\sum_{m=1}^{\infty}\sum_{b_1}\cdots\sum_{b_m}c(\alpha)_{b_1}\cdots c(\alpha)_{b_{m-1}}\left(d(\alpha)_{b_m}\;\theta_{a+b_1+\cdots+b_m,t+1-m}\right),$$

the probability distributions that the two specifications (with and without interactions)) generate on the observables of interest, $\{y_{a,t}\}_{a\in\mathbb{A},t\geq 1}$, are identical. Hence, one cannot identify from the stationary distribution of choices which specification generates the data. Similarly if T is finite, we showed in Section 2.1 that y_t is a well-defined linear function of y_0, θ_s and $\bar{\theta}$. Set $\theta_{at}'=y_{at}$. Then, the outcome probability distributions of an economy with habits α_1, interactions α_3 and shocks $(\theta_{a,t})$ and one with no habits, no interactions and shocks $(\theta_{a,t}')$ are identical. ■

An intuition about this result can be obtained by considering the infinite horizon case and loosely reducing the identification of dynamic conformity economies to the problem of distinguishing a VAR from an MA(∞) process. Stacking in a vector $\mathbf{y}_t$ (resp. θ_t) the actions $y_{a,t}$ over the index $a\in\mathbb{A}$ (resp. the preference shocks $\theta_{a,t}$), policy functions can be loosely written as a VAR:

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

where $E(\delta_t\delta_{t-\tau})=0$ for all $\tau>0$, and $\bar{\theta}$ is their identical expected values. Let $\mathbf{L}$ be a lag operator.[36] Under standard stationarity assumptions, the VAR has an MA (∞) representation

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

for a sequence $\Psi_1, \Psi_2 \dots$ such that $(I_A - \Phi\mathbf{L})\left(I_A + \Psi_1\mathbf{L} + \Psi_2\mathbf{L}^2 + \dots\right) = I_A$. The argument in the proof of Theorem 5 therefore amounts to picking

$$\mathbf{y}_t = \theta_t' = \delta_t + \Psi_1\delta_{t-1} + \Psi_2\delta_{t-2} + ...$$

3.1 Main Result

The intuition about Theorem 5 we provided suggests that identification might be obtained under restrictions on the correlation structure of the stochastic process $(\theta_{a,t})_{a\in\mathbb{A},t\geq 1}$. More specifically, assume that preference shocks are affected by a vector of observable time-varying individual characteristics, $x_{a,t}$, as follows:[37]

$$\theta_{a,t} = \gamma x_{a,t} + \delta(x_{a-1,t} + x_{a+1,t}) + u_{a,t}$$

where $u_{a,t}$ is an error term which is unobserved by the econometrician but observed by the agents and γ, δ are parameters to be estimated jointly with the deep preference parameters of the economy, $(\alpha_1, \alpha_3, \beta)$. The identification result will formally require $\gamma \neq 0$;[38] when also $\delta \neq 0$, the economy is characterized by *contextual peer effects*, that is, individual preference shocks are directly affected by friends' covariates. We make the following two classical assumptions on observables.

Assumption 2 (Exogeneity) $\forall s, t, E(u_s|x_t) = 0$.

Assumption 3 (Full rank) $\forall s, \forall a, \forall B \subset N$ *finite,* $x_{a,s}$ *does not depend linearly on* $\{x_{b,s}\}_{b\in B}$.

Assumption 2 requires that observables in any given period t be uncorrelated with unobservables in any period s. Assumption 3 requires that the matrix formed of elements in $\{x_{b,s}\}_{b\in B}$ for any finite subset B, has full rank. Furthermore, in the infinite horizon case, when equilibria are not necessarily unique, we maintain the selection that the equilibrium is the limit of equlibria of finite economies.

An important role in our analysis will be played by restrictions on the temporal memory of the $(\theta_{a,t})_{a\in\mathbb{A},t\geq 1}$ process. More specifically,

[36] $\mathbf{L}$ is a lag operator, i.e., $\mathbf{L}\delta_t := \delta_{t-1}$ for any period t. Polynomials of the lag operator can be used, follow similar rules of multiplication and division as do numbers and polynomials of variables, and this is a common notation for autoregressive moving average models. See e.g. Hamilton (1994) chapter 2 for an in-depth discussion of lag operators. See also Hamilton (1994) chapter 3 for a discussion of invertibility and equivalent representation arguments.

[37] Our analysis directly extends to the case of several characteristics.

[38] We however suggest that a related result can be obtained if $\gamma = 0$; see the discussion after the outline of the proof of Theorem 6, p. 24.

Definition 3 *Observables x have* finite temporal memory *if there exists some duration M such that $\forall a, b, E(x_{a,t}|x_{b,s}) = E(x_{a,t})$ if $t > s + M$.*

Under *finite temporal memory*, thus, correlation in observables across time does not extend for more than M periods.

We now state our main identification result.

Theorem 6 (Identification) *Consider a* dynamic linear conformity economy *such that the stochastic process $(\theta_{a,t})_{a\in\mathbb{A},t\geq 1}$ satisfies $\theta_{a,t} = \gamma\, x_{a,t} + \delta\,(x_{a-1,t} + x_{a+1,t}) + u_{a,t}$ under Assumptions 2 and 3. Suppose that $\alpha_1 \neq 0$ and $\gamma \neq 0$. If T is infinite and x has* finite temporal memory*, then the economy is identified. If T is finite, then the economy is identified even without restrictions on temporal memory.*

We provide an outline of the proof here and fill all details in the formal proof in the Appendix. It is useful to proceed in several steps to better illustrate the intuition of the results and, in particular, to stress the role that dynamic forward looking behaviour plays in the result.

No contextual effects. Consider first the special case in which $\delta = 0$ and individual preference shocks are not directly affected by friends' covariates:

$$\theta_{a,t} \;=\; \gamma\, x_{a,t} \,+\, u_{a,t}.$$

Observe that $E\left[\theta_{a+b,\tau}\middle|\, \theta^t\right] = \gamma E\left[x_{a+b,\tau}\middle|\, \theta^t\right] + E\left[u_{a+b,\tau}\middle|\, \theta^t\right] = \gamma E\left[x_{a+b,\tau}\middle|\, x^t\right] + E\left[u_{a+b,\tau}\middle|\, u^t\right]$ for $\tau > t$, since the agent observes both x and u. Moreover, $E\left[x_{a+b,\tau}\middle|\, x^t\right]$ is a function of x^t, which is known.

I) Infinite horizon. If $T = \infty$, our characterization of the equilibrium, in Theorem 1, can be re-written into the following econometric equation:

$$\begin{aligned} y_{a,t} \;&=\; \sum_{b\in\mathbb{A}} c_b\, y_{a+b,t-1} + \gamma \sum_{b\in\mathbb{A}} d_b\, x_{a+b,t} \\ &\quad + \gamma \sum_{\tau=t+1}^{\infty} \sum_{b\in\mathbb{A}} e_{b,\tau-t}\, E\left[x_{a+b,\tau}\middle|\, x^t\right] \,+\, \varepsilon_{a,t} \end{aligned}$$

where the error term $\varepsilon_{a,t}$ is a linear combination of own and others' current unobservables and expectations of future unobservables.

Identification then obtains if (i) we can consistently estimate "reduced-form" parameters c_b in the equation above; (ii) the mapping expressing c_b as functions of α_1, α_3 and β is injective, so that a unique vector of structural parameters can be obtained. We already showed that (ii) is true in Theorem 3-(i). As for (i), however, consistent estimates cannot be obtained by simple regression as past outcomes $(y_{a+b,t-1})_{b\in\mathbb{A}}$ are generally endogenous, for instance due to correlation between $\theta_{a,t}$ and $\theta_{a+b,t-1}$. We can nonetheless use an instrumental variables approach, with lagged observed shocks $x_{a+b,t-M-1}$ as instruments for past outcomes $y_{a+b,t-1}$.[39] This approach is valid under our assumptions. Indeed, Assumption 3 ensures that the instruments are not perfectly correlated; Assumption 2 and *finite temporal memory* together ensure that the appropriate exclusion restrictions are satisfied; specifically, finite memory ensures that the instruments are not correlated with the expectations of future covariates, $E\left[x_{a+b,\tau}\right| x^t\big]$. Finally, when $\gamma \neq 0$ and $\alpha_1 \neq 0$, these instruments are correlated with the endogenous regressors.

II) Finite horizon. If T is finite, our characterization of the equilibrium, in Theorem 1, can be re-written into the following econometric equation:

$$\begin{aligned} y_{a,t} \quad &= \quad \sum_{b\in\mathbb{A}} c_{b,T-(t-1)}\, y_{a+b,t-1} + \gamma \sum_{b\in\mathbb{A}} d_{b,T-(t-1)}\, x_{a+b,t} \\ &\quad + \gamma \sum_{\tau=t+1}^{T} \sum_{b\in\mathbb{A}} e_{b,T-(t-1),\tau-t}\, E\left[x_{a+b,\tau}\right| x^t\big] \,+\, \varepsilon_{a,t} \end{aligned}$$

The proof and the procedure adopted for proving identification for $T=\infty$ in *I)* also works when T is finite, as long as $M < T$. However, the restriction on x's correlation structure is actually not needed to guarantee identification in this case, because we can exploit the lack of stationarity of the dynamic equilibrium in a fundamental manner. Indeed, consider decisions at T and $T-1$. In the last period, the first-order condition of agent a's optimization problem admits the following simple expression:

$$y_{a,T} \quad = \quad \alpha_1\, y_{a,T-1} + \alpha_3\,(y_{a-1,T} + y_{a+1,T}) + \alpha_2\,\gamma\, x_{a,T} + \varepsilon_{a,T}, \tag{9}$$

where $\varepsilon_{a,T} = \alpha_2\, u_{a,T}$. This econometric equation expresses own current outcome as a function of two endogenous variables: own past outcome and friends' current outcomes. Under Assumptions 2 and 3, this equation can be consistently estimated by using as instruments, for instance, own past observed shock $x_{a,T-1}$ and friends' current observed shocks $(x_{a-1,T}+x_{a+1,T})$. These instrumental regressions provide direct estimates of α_1 and α_3.

To identify β, however, we need to focus on equilibrium outcomes at $T-1$. We show in Appendix E.2 that the first-order condition at $T-1$ is equivalent to the following econometric

[39] We adopt an asymptotic framework based on an arbitrarily large number of replications of the economy. Thus, outcomes and errors should in principle be indexed by replication r in the econometric equations. We omit these additional indices for clarity.

equation:

$$
\begin{aligned}
[1+\beta\,\alpha_1\,(1-c_{1,1})]\,y_{a,T-1} \;=\;& \alpha_1\,y_{a,T-2}+\alpha_3\,(y_{a-1,T-1}+y_{a+1,T-1})\\
&+\beta\,(\alpha_1-c_{1,1}(1-2\alpha_3))\,y_{a,T}\\
&+\gamma\,\alpha_2\,x_{a,T-1}+\gamma\,\beta\,\alpha_2\,c_{1,1}\,x_{a,T}+\varepsilon_{a,T-1}
\end{aligned}
\tag{10}
$$

where $c_{1,1}$ is independent of β and $\varepsilon_{a,T-1}$ includes preference shocks as well as differences between expected and realized outcomes at T. When $\beta = 0$, agents are myopic and first-order conditions at $T-1$ and T have the same form. In contrast, when $\beta > 0$, forward looking decisions taken in $T-1$ depend on their anticipated impact on the agent's utility at T. As a consequence, the econometric equation expresses current outcome of agent a (at $T-1$) as a linear function of own future outcome $y_{a,T}$, besides own past outcome $y_{a,T-2}$ and friends' current outcomes $(y_{a-1,T-1}+y_{a+1,T-1})$. Three valid instruments are enough to provide consistent estimates. When $\alpha_3 \neq 0$, these instruments could be, for instance, $x_{a,T-2}$, $(x_{a-1,T-1}+x_{a+1,T-1})$, and $(x_{a-2,T-1}+x_{a+2,T-1})$; that is, own past observed shocks, friends' current observed shocks and friends of friends' current observed shocks.

Contextual effects. Consider first the case in which T is infinite. Equilibrium characterization can now be written as follows:

$$
\begin{aligned}
y_{a,t} \;=\;& \sum_{b\in\mathbb{A}} c_b y_{a+b,t-1}+\gamma\sum_{b\in\mathbb{A}} d_b x_{a+b,t}+\delta\sum_{b\in\mathbb{A}} d_b(x_{a-1+b,t}+x_{a+1+b,t})+\gamma\sum_{\tau=t+1}^{\infty}\sum_{b\in\mathbb{A}} e_{b,\tau-t}E\left[x_{a+b,\tau}|x^t\right]\\
&+\;\delta\sum_{\tau=t+1}^{\infty}\sum_{b\in\mathbb{A}} e_{b,\tau-t}(E\left[x_{a-1+b,\tau}|x^t\right]+E\left[x_{a+1+b,\tau}|x^t\right])+\;\varepsilon_{a,t}.
\end{aligned}
$$

Under finite temporal memory, the argument exploited in case *I)* above, with no contextual effects, directly applies. We can still use lagged observed shocks $x_{a+b,t-M-1}$ as instruments for past outcomes $y_{a+b,t-1}$.

Consider next the case in which T is finite. The econometric expression of the first-order condition at T is:

$$
y_{a,T} \;=\; \alpha_1 y_{a,T-1}+\alpha_3(y_{a-1,T}+y_{a+1,T})+\alpha_2\gamma x_{a,T}+\alpha_2\delta(x_{a-1,T}+x_{a+1,T})+\varepsilon_{a,T}
$$

Contextual peer effects prevent, of course, the use of current observed shocks of friends as instruments. However, and related to the approach developed in Bramoullé et al. (2009), the current observed shocks of friends of friends, $(x_{a-2,T}+x_{a+2,T})$, can now be used as an instruments for friends' current outcome.

At $T-1$ we now have:

$$
\begin{aligned}
[1+\beta\alpha_1(1-c_{1,1})]y_{a,T-1} \;=\;& \alpha_1 y_{a,T-2}+\alpha_3(y_{a-1,T-1}+y_{a+1,T-1})+\beta(\alpha_1-c_{1,1}(1-2\alpha_3))y_{a,T}\\
&+\gamma\alpha_2 x_{a,T-1}+\delta\alpha_2(x_{a-1,T-1}+x_{a+1,T-1})\\
&+\gamma\beta\alpha_2 c_{1,2}x_{a,T}+\delta\beta\alpha_2 c_{1,1}(x_{a-1,T}+x_{a+1,T})+\varepsilon_{a,T-1}
\end{aligned}
$$

And we can now, for instance, use own past observed shocks, $x_{a,T-2}$, friends' past observed shocks, $(x_{a-1,T-2} + x_{a+1,T-2})$, and friends of friends' past observed shocks, $(x_{a-2,T-2} + x_{a+2,T-2})$, as instruments.

This outline of the proof of Theorem 6 shows that the fundamental effect of forward-looking behaviour at equilibrium consists in having current actions partly determined by expectations over future actions. Because of the general pattern on correlation over time and space of the stochastic process for the shocks, the future actions of any agent depend, in turn, on all the agents' future shocks. This is the case also with no contextual effects. In the econometric equation for the infinite horizon economy, with no contextual effects, expectations over future actions are solved out for in terms of expected future observables of all the agents, collected in the term $\sum_{\tau=t+1}^{\infty} \sum_{b \in \mathbb{A}} e_{b,\tau-t} E\left[x_{a+b,\tau} | x^t\right]$. While the dependence of current actions on expectations over future actions in the dynamic economy complicates identification, under *finite temporal memory*, observables sufficiently in the past do not affect future shocks (and hence the expectations of future actions) and hence they can be used as valid instruments for identification.

While the same issue appears in the finite horizon economy, in this case we can exploit the non-stationarity of forward-looking behaviour in equilibrium to identify the parameters of the model. In particular, the reduced form parameters in the econometric equations, which determine any agent's actions, depend on t (more precisely, on the number of periods to the end of the economy T). The properties of this dependence can be exploited to identify the model. As we noted in the outline of the proof, when $\alpha_3 \neq 0$, we can exploit, for instance, either friends of friends' current observed shocks $(x_{a-2,T-1} + x_{a+2,T-1})$ or friends' past observed shocks $(x_{a-1,T-2} + x_{a+1,T-2})$. In other words, valid instruments can be found back in time and further out in social space: thanks to the dynamic nature of the model, we can make use of the two dimensions of time and space to help correct the endogeneity problems arising in the regressions. We can combine social and temporal lags in ways unfeasible in static models of peer effects or in individual dynamic models.

Importantly, this whole discussion about the role of forward-looking behaviour at equilibrium and non-stationarity in identification naturally extends to the introduction of contextual peer effects. Indeed, in the stationary infinite horizon economy contextual effects do not require any additional instrument for identification. In the non-stationary economy instead, we need to rely either on friends of friends' past observed shocks, $(x_{a-2,T-2} + x_{a+2,T-2})$, or on friends' observed shocks further in the past, $(x_{a-1,T-3} + x_{a+1,T-3})$. Once again, we can make use of the two dimensions of time and space to strengthen identification.

Identification may also hold, under additional assumptions on their correlation structure, even when preference shocks $\theta_{a,t}$ do not depend on exogenous covariates, that is, when $\gamma = 0$. For instance, suppose that the stochastic process is of expectation zero and has finite temporal memory. Then, adapting classical techniques from dynamic panels (see e.g., Arellano and Bond (1991)), own and friends' lagged outcomes, $y_{a,t-M-1}$ and $(y_{a-1,t-M-1} + y_{a+1,t-M-1})$, can be used as instruments for own and friends' current outcome, $y_{a,t}$ and $(y_{a-1,t} + y_{a+1,t})$. Unlike with dynamic panels, however, the interplay of habits and social interactions means that own current outcome may be affected by own lagged outcome, friends' lagged outcome, friends of friends' lagged outcomes, etcetera. Thus the strength of identification generally depends on α_1, α_3 and β.

It is also apparent from the outline of the proof of Theorem 6 that we can extend our previous arguments to the case in which we do not observe outcomes $\mathbf{y}_t$ and observables $\mathbf{x}_t$ in all periods t. Indeed, if observables display finite temporal memory of length $M < T$, identification holds under the same conditions as in Theorem 6 if we observe outcomes at any t, $t-1$, $t-2$ and observables from $t - M - 1$ to t.[40]

Finally, we note that our identification strategy depends on the ability to exploit exogenous shocks on the preference parameter: applying our strategy to some observable x which is not exogenous would invalidate it and could e.g., generate positive estimates of social interactions even if $\alpha_3 = 0$, whenever unobservables are correlated with x and also across time and space. This is related to the standard problem of endogeneity which occurs in peer effects estimations in static models. Of course, as we show, standard solutions of this problem in the applied literature, involving random groups or exogenous shocks, still fail when they do not account for dynamic effects and these dynamic effects are present.

Remark (Myopic economy - continued) We can now turn to the question of whether behaviour in the dynamic linear conformity economies, with rational forward looking agents, can be identified from behaviour in the same economies with myopic agents introduced in Section 2.1. As we already noticed, the coefficients of the policy function in economies with myopic agents are equal to the ones of the policy function of a dynamic linear conformity economy with either $\beta = 0$ or at $t = T$. In this sense, myopic models are nested within the class of dynamic models we study and forward looking and myopic behaviour can be separately identified as an immediate consequence of Theorem 6.

Furthermore, and most importantly, forward looking and myopic behaviour can be identified also when only outcomes at any arbitrary time t ($< T$ if the economy has finite horizon) are observed. This is a direct consequence of Theorem 2-(ii-iii). There we show in fact that, when $\beta = 0$, and/or in the last period T of a finite horizon economy, the coefficients of the policy function form an exponentially declining sequence in $|\,b\,|$. When instead $\beta > 0$ and the economy has an infinite horizon, or when $\beta > 0$ and $t < T$ in a finite horizon economy, the coefficients of the policy function do not decline exponentially. It follows therefore that the policy function coefficients in a myopic economy decline exponentially in $|\,b\,|$. Showing that the estimated sequence $(c_{b,l})$ is

[40] If, furthermore, observables display no temporal correlation, that is, if $\forall a, b$ and $t > s$, $E(x_{a,t}|x_{b,s}) = E(x_{a,t})$, then conditional expectations $E\left[x_{a+b,\tau}|x^t\right]$ are equal to zero in the econometric equations and the economy is fully identified from observations of outcomes and observables of *any* consecutive periods t, $t-1$, $t-2$ only.

not exponential in $|b|$, where $l = T - (t-1) > 1$, implies that these data could not have been generated by a myopic model.[41]

Following this analysis, we can also sign the error we would make if we fitted a myopic (or a static) model to data generated by a dynamic model with forward looking agents. In fact, in Theorem 2-(iii) we show that, in a dynamic model with forward looking agents, at any time $t < T$, the sequence $(c_{b,l})$ declines faster than exponentially in $|b|$. As a consequence, erroneously fitting a myopic (or a static) model induces under-estimating $(c_{b,l})$ for b's close enough to a and overestimating them for all the more distant b's in social space.[42] While theoretically this could lead to both under- or over-estimating the social interaction parameter α_3, the Monte Carlo simulations we report in the next section suggest that erroneously fitting a static model generally leads to an under-estimation of the true value of this parameter. Thus, current estimates of social interactions may potentially be seriously biased downwards in contexts where forward-looking behaviour has a relevant effect.

Remark (Extensions) The identification strategy developed in this chapter can be adapted and extended to cover the more general economies we introduce in Section 2.3. We briefly outline the arguments here. In economies with *general network structures*, the policy functions are expressed as in equation (7). With a general (arbitrary) network structure, b cannot be interpreted any more as a sufficient measure of the social distance between an arbitrary agent a and $a+b$. As a consequence policy functions lose symmetry: the coefficient of the policy of a with respect to e.g., the past action of agent $a+b$ is $c_{a,b}$; that is, it depends on a as well as on b. Nonetheless, the coefficients are known functions of the network and the structural parameters. Under finite temporal memory and an infinite or sufficiently long finite time horizon, sufficiently lagged shocks continue to represent valid instruments for past actions. If anything, the (known) network structure should help with identification. In the economy with *social accumulation of habits*, the policy functions, in equation (8), depend on the accumulated stock $r_{a+b,t-1}$ in place of action $y_{a+b,t-1}$. But the idea of using lagged past shocks as instruments extends directly: we can simply instrument $r_{a+b,t-1}$ by $x_{a+b,t-M-1}$. More specifically, in finite horizon economies, non-stationary patterns can also be exploited for identification: for instance, social interactions might lose significance in the last periods in contrast to what predicted by habit models (or by simple behavioural imitation models, which we do not explicitly discuss).[43] Finally, in the economy with *global interactions*, provided that the average action $\rho(y)$ is well defined, the structure of the policy functions is maintained unchanged with respect to the the global interactions and α_4 is

[41] Equivalently, we can estimate the sequence $(\gamma d_{b,l})$ by simply regressing $y_{a,t}$ on $(x_{b,t})_{b\in\mathbb{A}}$ and then by showing that the estimated sequence is not exponential in $|b|$, where $l = T-(t-1) > 1$. The value of γ does not affect the rate of decline of the sequence $(\gamma d_{b,l})$, as long as $\gamma \neq 0$, and hence it can be disregarded in this argument.

[42] The same holds for the sequence $(d_{b,l})$.

[43] We thank an anonymous referee for this point.

absorbed in the coefficient (c_b, d_b, e_b), in equation (9). Then, after showing these new coefficients are injective, we can straighforwardly apply our identification strategy.

3.2 Monte Carlo Simulations

The strength of identification will generally be a function of the three structural parameters, capturing the complex interactions between habits, social interactions and discounting of future utilities, in connecting, in equilibrium, the actions across time. In this section, we explore this issue through Monte Carlo simulations where we evaluate the performance of Generalized Method of Moments (GMM) estimators of the parameters of interest obtained by exploiting moment conditions induced by dynamic economic equilibrium restrictions.

We concentrate on a three period $(T, T-1, T-2)$ conformity economy. We restrict the stochastic process $(\theta_{a,t})_{a\in\mathbb{A}, t\in\{T-2,T-1,T\}}$ by assuming no contextual effects, that is $\delta = 0$; therefore, $\theta_{a,t} = \gamma x_{a,t} + u_{a,t}$. Also, we assume $x_{a,t} \sim N(0,1)$ and $u_{a,t} \sim N(0,1)$, for $a \in \mathbb{A}$ and $t = T-2, T-1, T$.

The parameters we aim at estimating are $(\alpha_1, \alpha_3, \beta, \gamma)$. To this end, (i) we generate data from the stochastic process $(\theta_{a,t})_{a\in\mathbb{A}, t\in\{T-2,T-1,T\}}$ and (ii) for the dependent choice variables $y_{a,t}$, $t = T-2, T-1, T$, by using the unique equilibrium policy function in Theorem 1-(i).[44] In the benchmark specification whose output we present in Figures 4 and 5, we use $\alpha_1 = 0.2, \alpha_3 = 0.3, \alpha_2 = 0.2, \beta = 0.95, \gamma = 1.0, m = 1000$, and $B = 500$, where B is the Monte Carlo size, although we rely on multiple other designs for our evaluations.[45] We then apply the GMM procedure to the data we generated, using 2 moment conditions imposed by equations (9) for $t = T$:

$$E\left[\varepsilon_{a,T} \mid x_{a,T-1}\right] = 0 \tag{11}$$

$$E\left[\varepsilon_{a,T} \mid (x_{a-1,T} + x_{a+1,T})\right] = 0; \tag{12}$$

and 3 moment conditions from equation (10) for $t = T-1$:

$$E\left[\varepsilon_{a,T-1} \mid x_{a,T-2}\right] = 0 \tag{13}$$

$$E\left[\varepsilon_{a,T-1} \mid (x_{a-1,T-1} + x_{a+1,T-1})\right] = 0 \tag{14}$$

$$E\left[\varepsilon_{a,T-1} \mid (x_{a-2,T-1} + x_{a+2,T-1})\right] = 0. \tag{15}$$

[44] In order to make sure that the results are not influenced by initial observations $(y_{a,T-2})_{a\in\mathbb{A}}$, we used multiple designs with buffer periods of size 0, 10, 20, and 50. We either set the first buffer period values to zero or picked them iid from $N(0,1)$. For the rest of the buffer periods, we generated $x_{a,t} \sim N(0,1)$ and $u_{a,t} \sim N(0,1)$, and let the $y_{a,t}$ be determined by the equilibrium policy function in Theorem 1-(i). We obtained practically identical results in each, to the ones in Figures 4 and 5.

[45] Precisely, we have used in our Monte Carlo experiments the following designs: $N = 2m+1$, with $m = 10, 25, 50, 100, 500, 1000$, and 5000 agents on both sides of agent zero; $\alpha_3 = 0, .05, .1, .15, .20, .25, .30, .35, .40, .45, .50$; $\alpha_1 = p(1-2\alpha_3)$, and $\alpha_2 = (1-p)(1-2\alpha_3)$, where $p = 0, .1, .2, .3, .4, .5, .6, .7, .8, .9, 1$.

Given these 5 population moment conditions for each agent $a \in \mathbb{A}$, nonlinear in the parameters of interest, $(\alpha_1, \alpha_3, \beta, \gamma)$, the model is *overidentified.* Therefore it is not possible to solve the system of analogous sample moment conditions for a unique value of the parameter vector. Instead, the structural coefficients can be estimated using GMM.[46]

The results of our Monte Carlo simulations reported in Figure 4 are encouraging regarding the strength of identification: all coefficient estimates converge to the true values and standard errors shrink, as the number of agents becomes arbitrarily large. Interestingly, however, the convergence speed for the β estimate is much slower than for (α_1, α_3). We suggest this is a consequence of the fact that the GMM estimate of β is obtained exclusively off of equations (13-15), whereas (α_1, α_3) enter both T and $T-1$ period first order conditions, equations (13-15) and (11-12), respectively.

Our Monte Carlo allows us to study more in detail the issue of distinguishing myopic (or static) and dynamic economies. In particular we aim at studying the error associated with estimates of α_1 and α_3 that ignore the dynamic structure of the economy when the true model is a dynamic linear conformity economy. To this end, we perform the following experiment: For the same data drawn in each Monte-Carlo round, we estimate (i) the full GMM (5 moments-2 periods), (ii) GMM with $T-1$-period data only and 3 moment conditions for period $T-1$, and finally (iii) myopic (static) GMM, with $T-1$-period data where the econometrician (mistakenly) believes that it is T-period data and using the 2 moment conditions for period T. Figure 5 reports the outcome of the experiment: the full GMM in red, the GMM with $T-1$-period data only in blue, and the myopic (static) GMM in green.

As it is apparent from the Figure, both the full GMM and the GMM with $T-1$-period data consistently estimate α_1 and α_3, although standard errors under the full GMM are much smaller, due to the fact that this procedure uses more information. The estimates of both parameters under myopic (static) GMM are instead biased, evidencing the mis-specification of the true dynamic forward-looking equilibrium data generating process. Most importantly, while theoretically we can only sign the error in terms of the patterns of $(c_{b,l})$ and $(d_{b,l})$, in the simulations we can directly sign the bias in the estimates: the myopic (static) GMM under-estimates both α_1 and α_3 with respect to the true value; that is, it over-estimates the effect of agents' own shock in their preferences, α_2. Thus, an econometrician who ignores the true dynamic structure with forward-looking behaviour might potentially end up obtaining estimates of social interactions that are biased downwards.

4 Conclusion

Social interactions provide a rationale for several important phenomena at the intersection of economics and sociology. As we noted in the Introduction, however, the theoretical and empir-

[46] We follow closely the setup of Section 6.5 of Cameron and Trivedi (2005) on nonlinear instrumental variables.

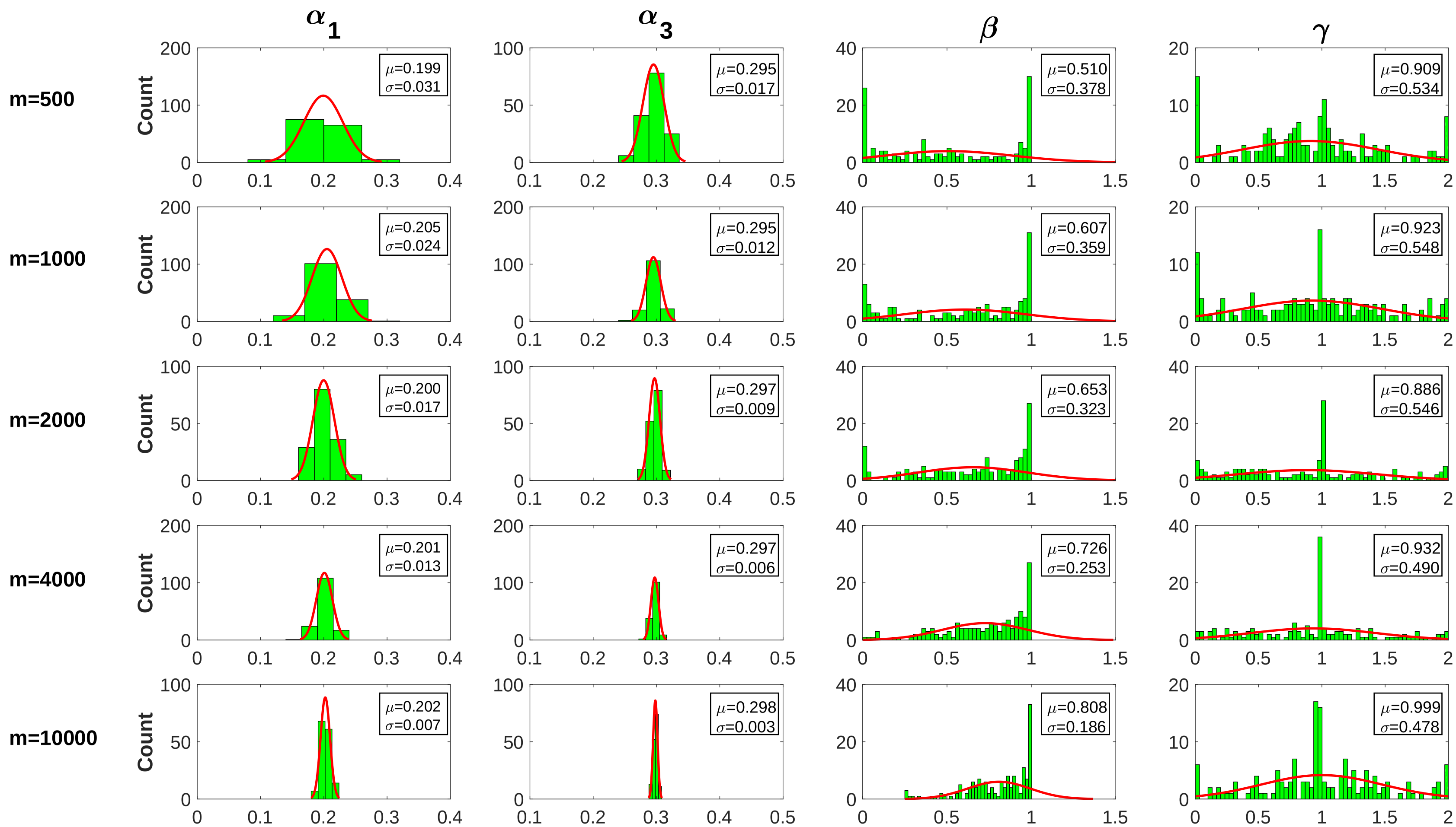

Figure 4: Convergence of the estimates to the true values of $(\alpha_1, \alpha_3, \beta, \gamma) = (0.2, 0.3, 0.95, 1)$.

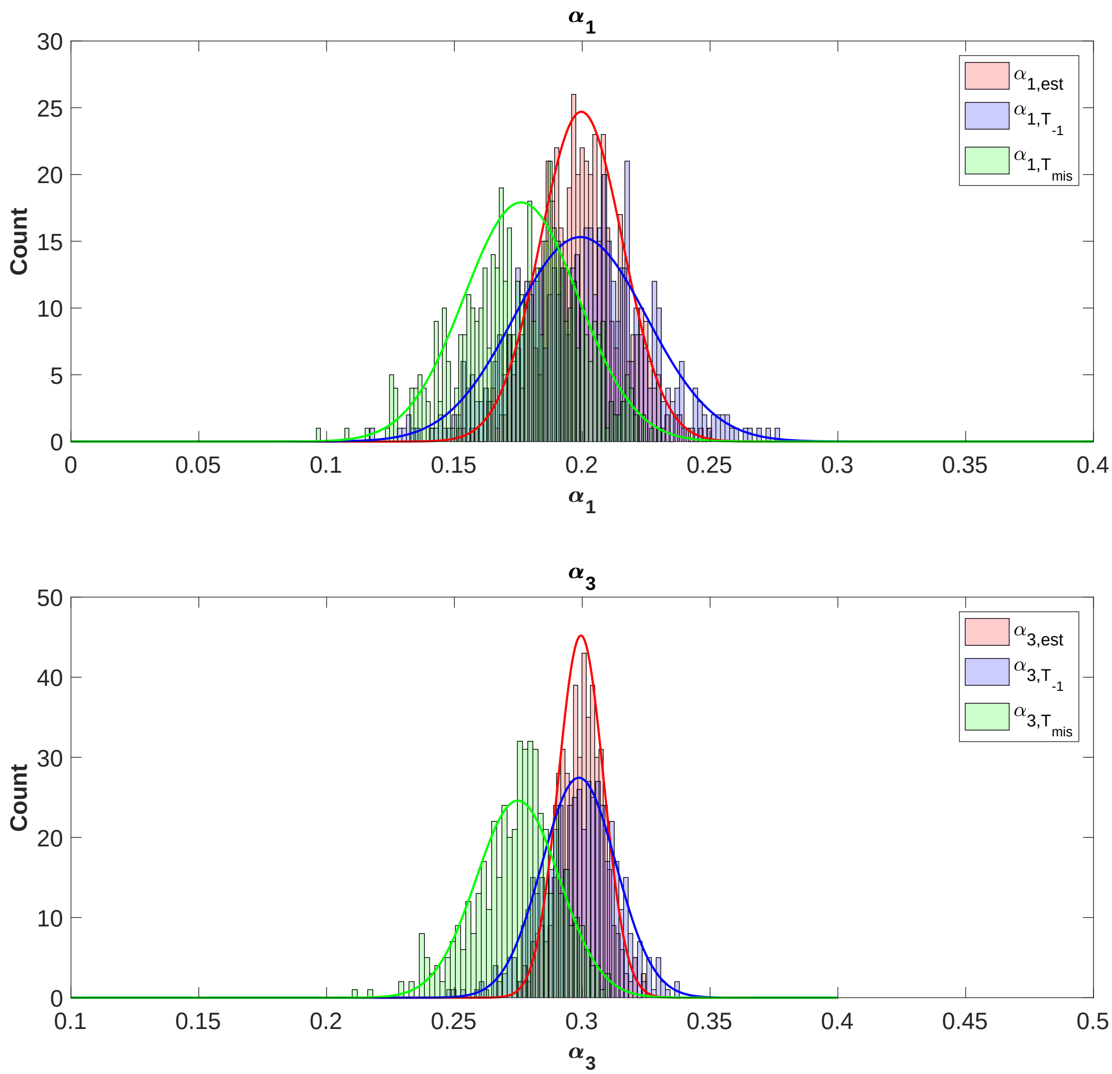

Figure 5: Distribution of the estimates $\hat{\alpha}_1$ and $\hat{\alpha}_3$ for the true values of $(\alpha_1, \alpha_3, \beta, \gamma) = (0.2, 0.3, 0.95, 1)$, under three different specifications.

ical study of economies with social interactions has been hindered by both mathematical and conceptual problems.

In this paper we show how some of these obstacles to the study of economies with social interactions can be overcome. Admittedly, we restrict our analysis to linear economies, but in this context we are able to prove several desirable theoretical and computational properties of equilibria and to exploit the properties of dynamic equilibria we characterize to produce positive identification results both in stationary and non-stationary economies.

We conclude that the class of dynamic linear economies with social interactions we have studied in this paper can be fruitfully employed in applied and empirical work.

References

Aguirregabiria, V. and Nevo, A. (2013), "Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games," *Advances in Economics and Econometrics. Tenth World Congress.* Volume 3, Econometrics. (May 2013). D. Acemoglu, M. Arellano, and E. Deckel, eds.

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

Arellano, M. and Bond, S. (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, **58**, 277-297.

Bajari, P. Hong, H. and Nekipelov, D. (2013), "Econometrics for Game Theory," *Advances in Economics and Econometrics. Tenth World Congress.* Volume 3, Econometrics. (May 2013). D. Acemoglu, M. Arellano, and E. Deckel, eds.

Bajari, P., L. Benkard, and J. Levin (2007), "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75, 133170.

Ballester, C., Calvó-Armengol, A., Zenou, Y. (2006), "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, **74**, (5), 1403-1417.

Bartle, R. G. (1976), *The Elements of Real Analysis*, second edition. John Wiley & Sons, New York.

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.

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.

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.

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., Brock, W., Durlauf, S., and Jayaraman, R. (2015), "Linear Social Interactions Models," *Journal of Political Economy*, **123**, 2, 444-496.

Blume, L. and Durlauf, S., (2001), "The Interactions-Based Approach to Socioeconomic Behaviour," 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.

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

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.

Cameron, A. C. and Trivedi, P. K. (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press, Cambridge.

Carmona, A., Encinas, A. M., Gago, S., Jiménez, M. J. and Mitjana, M. (2015), "The Inverses of Some Circulant Matrices," In *Appl. Math. Comput*, **270**, 785-793.

Cole, H. L., Mailath, G. J. and Postlewaite, A. (1992), "Social Norms, Savings Behaviour, and Growth," *Journal of Political Economy*, **100**, 1092-1125.

Cooke, R. G. (1950), *Infinite Matrices and Sequence Spaces*, London: Macmillan.

Davezies, L., d'Haultfoeuille, X., and Fougère, D., (2009), "Identification of Peer Effects Using Group Size Variation," *The Econometrics Journal*, **12**, 3, 397-413.

Davis, P. J. (1970), *Circulant Matrices*, New York: Wiley.

De Paula, A. (2009), "Inference in a Synchronization Game with Social Interactions," *Journal of Econometrics*, **148**, 5671.

Duffie, D., Geanakoplos, J., Mas-Colell, A., and McLennan, A. (1994), "Stationary Markov Equilibria," *Econometrica*, **62**, 745-781.

Dunford, N. and Schwartz, J. T. (1958), *Linear Operators. Part I. General Theory*, New York: John Wiley and Sons.

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. 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 Behaviour," *Journal of Political Economy*, **100**, 51, 966-991.

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

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.

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.

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.

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.

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

Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1029-1054.

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

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

Ioannides, Y. M. and Soetevent, A. R. (2007), "Social Networking and Individual Outcomes Beyond the Mean Field Case," *Journal of Economic Behaviour and Organization*, **64**, 369-390.

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

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

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

Liggett, T.M. (1985) *Interacting Particle Systems*, Springer Verlag, Berlin.

Magnac, T. and Thesmar, D. (2002), "Identifying Dynamic Discrete Decision Processes," *Econometrica*, **70** (2), 801-816.

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.

Mertens, J.-F. (2002), "Stochastic Games," In R.J. Aumann and S. Hart, eds., *Handbook of Game Theory*, Vol. **3** (North-Holland,Amsterdam), Chapter 47, 1809-1832.

Mertens, J.-F. and Parthasarathy, T. (1987), "Equilibria for Discounted Stochastic Games," Research paper 8750, CORE, Université Catholique de Louvain.

Moody, J., (2001), "Race, School Integration, and Friendship Segregation in America", *American Journal of Sociology*, **107**, 3, 679-716.

Nowak, A. S., (2003), "N-Person Stochastic Games: Extensions of the Finite State Space Case and Correlation", *Stochastic Games and Applications*, Ed. Abraham Neyman and Sylvain Sorin, 93-106.

Özgür, O., (2010), "Local Interactions", In J. Benhabib, M. O. Jackson, and A. Bisin, eds., *Handbook of Social Economics*, Vol. **1A** (North-Holland, Amsterdam), Chapter 13, 587-644.

Peski, M. (2007), "Complementarities, Group Formation, and Preferences for Similarity," mimeo.

Pettofrezzo, A. J and Byrkit, D. R. (1970), *Elements of Number Theory*, Englewood Cliffs: Prentice Hall.

Rozen, K. (2010), "Foundations of Intrinsic Habit Formation," *Econometrica*, 78 (4), 13411373.

Rust, J. (1994), "Structural Estimation of Markov Decision Processes," In R. Engle and D. McFadden, eds., *Handbook of Econometrics*, Vol. **4** (North-Holland,Amsterdam), Chapter 51, 3081-3143.

Santos, M. S. (1991), "Smoothness of the Policy Function in Discrete Time Economic Models," *Econometrica*, **59**, 1365-82.

Vieille, N. (2002), "Stochastic Games: Recent Results," In R.J. Aumann and S. Hart, eds., *Handbook of Game Theory*, Vol. **3** (North-Holland,Amsterdam), Chapter 48, 1833-1850.

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

APPENDIX

This Appendix contains proofs of the major results in the main text. There are a total of seven lemmas these proofs depend upon. As a general rule, we have stated the technical lemmas we need when we need them, and relegated the detailed proofs to the Technical Appendix (not for publication), to make the article flow better. All notation is as defined in the main text unless explicitly noted otherwise.

A Proof of Theorem 1: Existence and Uniqueness

By exploiting the linearity of the policy functions, our method of proof is constructive, producing also a direct and useful recursive computational characterization for the parameters of the symmetric policy function at equilibrium. We implement the proof in three main steps, by induction on the length of the economy.

Step 1: Existence, uniqueness and the convex form for $T = 1$. One can interpret this step either as a one-period economy or the last period of a finite-horizon economy, by considering initial histories of any finite length s. Let any history $(y^0, \theta^1) = (y_{-(s-1)}, \theta_{-(s-2)}, \ldots, y_{-1}, \theta_0, y_0, \theta_1)$ of previous choices and preference shock realizations be given. Agent a solves

$$\max_{y_{a,1} \in Y} \left\{ -\alpha_1 \left(y_{a,0} - y_{a,1}\right)^2 - \alpha_2 \left(\theta_{a,1} - y_{a,1}\right)^2 - \alpha_3 \left(y_{a-1,1} - y_{a,1}\right)^2 - \alpha_3 \left(y_{a+1,1} - y_{a,1}\right)^2 \right\} \quad \text{(A.1)}$$

The first order condition

$$2 \left[\alpha_1 \left(y_{a,0} - y_{a,1}\right) + \alpha_2 \left(\theta_{a,1} - y_{a,1}\right) + \alpha_3 \left(y_{a-1,1} - y_{a,1}\right) + \alpha_3 \left(y_{a+1,1} - y_{a,1}\right)\right] = 0$$

implies that

$$y_{a,1} = \Delta_1^{-1} \left(\alpha_1\, y_{a,0} + \alpha_2\, \theta_{a,1} + \alpha_3\, y_{a-1,1} + \alpha_3\, y_{a+1,1}\right) \quad \text{(A.2)}$$

where $\Delta_1 := \alpha_1 + \alpha_2 + 2\alpha_3 > 0$. This choice is feasible (in Y) since it is a convex combination of elements of Y, a convex set by assumption. The objective function (A.1) is strictly concave in $y_{a,1}$, thus $y_{a,1}$ in (A.2) is the unique optimizer. The form in (A.2) maps bounded measurable policy functions f_{a-1} and f_{a+1}, for $a-1$ and $a+1$ respectively, into a bounded measurable function for a. The system of such maps, one for each agent a, defines, given any s-length history (y^0, θ^1), an operator $L_1 : \mathbf{B} \to \mathbf{B}$ that acts on the family of bounded measurable functions $f = (f_a) \in \mathbf{B} := \prod_{a \in \mathbb{A}} B\left((Y \times \Theta)^s, Y\right)$ according to

$$\left(L_1 f\right)_a \left(y^{t-1}, \theta^t\right) = \Delta_1^{-1} \left(\alpha_1\, y_{a,0} + \alpha_2\, \theta_{a,1} + \alpha_3\, f_{a-1}\left(y^{t-1}, \theta^t\right) + \alpha_3\, f_{a+1}\left(y^{t-1}, \theta^t\right)\right)$$

Clearly, L_1 is a self-map. We show next that it is a contraction. Endow $B\left((Y \times \Theta)^s, Y\right)$ and $\mathbf{B}$ with the sup norm $\|\cdot\|_\infty$ which makes them into Banach spaces. Pick $f, f' \in \mathbf{B}$. For all $\left(y^0, \theta^1\right)$

$$\begin{aligned}
\Big| \left(L_1 f\right)_a \left(y^0,\theta^1\right) - \left(L_1 f'\right)_a \left(y^0,\theta^1\right) \Big| \;&=\; \Delta_1^{-1} \Big| \alpha_1\, y_{a,0} + \alpha_2\, \theta_{a,1} + \alpha_3\, f_{a-1}\left(y^0,\theta^1\right) + \alpha_3\, f_{a+1}\left(y^0,\theta^1\right) \\
&\quad - \alpha_1\, y_{a,0} - \alpha_2\, \theta_{a,1} - \alpha_3\, f'_{a-1}\left(y^0,\theta^1\right) - \alpha_3\, f'_{a+1}\left(y^0,\theta^1\right) \Big| \\
&= \Delta_1^{-1} \Big| \alpha_3\, \left(f_{a-1}\left(y^0,\theta^1\right) - f'_{a-1}\left(y^0,\theta^1\right)\right) \\
&\quad + \alpha_3\, \left(f_{a+1}\left(y^0,\theta^1\right) - f'_{a+1}\left(y^0,\theta^1\right)\right) \Big| \\
&\leq \left(\frac{\alpha_3}{\Delta_1}\right) \Big| f_{a-1}\left(y^0,\theta^1\right) - f'_{a-1}\left(y^0,\theta^1\right) \Big| \\
&\quad + \left(\frac{\alpha_3}{\Delta_1}\right) \Big| f_{a+1}\left(y^0,\theta^1\right) - f'_{a+1}\left(y^0,\theta^1\right) \Big| \\
&\leq \left(\frac{\alpha_3}{\Delta_1}\right) \left(\left\|f_{a-1} - f'_{a-1}\right\|_\infty + \left\|f_{a+1} - f'_{a+1}\right\|_\infty\right) \\
&\leq \left(\frac{2\alpha_3}{\Delta_1}\right) \|f - f'\|_\infty
\end{aligned}$$

Since the above inequality holds for each a, then

$$\|L_1 f - L_1 f'\|_\infty \;\leq\; \left(\frac{2\alpha_3}{\Delta_1}\right) \|f - f'\|_\infty$$

The coefficient $2\alpha_3\, \Delta_1^{-1} < 1$ since $\alpha_1 + \alpha_2 > 0$ thanks to Assumption 1-2.. Hence L_1 is a contraction mapping on $\mathbf{B}$. Thus, by Banach Fixed Point Theorem (see e.g., Aliprantis and Border (2006), p.95) L_1 has a unique fixed point, i.e., a unique family $f^* = (f_a^*)$ in $\mathbf{B}$ that satisfies the system of first-order conditions for all agents. Consider now $\mathbf{B}_s$ the subset of $\mathbf{B}$ that includes symmetric families of bounded measurable maps in the sense that for any $a, b \in \mathbb{A}$ and any given history (y^0, θ^1), $f_{a+b}(y^0, \theta^1) = f_a(R^b\, y^0, R^b\, \theta^1)$, where R^b is the canonical shift operator in the direction b. It is easy to show that $\mathbf{B}_s$ is a closed subset of $\mathbf{B}$ and that L_1 maps $\mathbf{B}_s$ into itself. Since L_1 is a contraction mapping, its unique fixed point then lies necessarily in $\mathbf{B}_s$. This establishes that the unique solution needs to be symmetric. Furthermore, as we show in the next Lemma, by applying the operator L_1 directly to policy coefficient sequences, this unique symmetric family takes the convex combination form as in the statement of Theorem 1. Let

$$G := \left\{ \begin{array}{l}
g : Y \times \Theta^s \to Y \text{ s.t.} \\
g(y_0,\ \theta^1) = \sum_{a\in\mathbb{A}} c_a\, y_{a,0} + \sum_{a\in\mathbb{A}} d_a\, \theta_{a,1} + \sum_{\tau=t+1}^{T} \sum_{a\in\mathbb{A}} e_{a,\tau-t}\, E\left[\theta_{a,\tau}|\theta^1\right] \\
\text{with} \\
\;\text{(i) } c_a, d_a, e_a \geq 0 \text{ and } \sum_{a\in\mathbb{A}} \left(c_a + d_a + \sum_{\tau=t+1}^{T} e_{a,\tau-t}\right) = 1 \\
\;\text{(ii) } (\tfrac{1}{2})c^{a+1} + (\tfrac{1}{2})c^{a-1} \geq c^a, \forall a \neq 0 \\
\;\text{(iii) } c^b \leq c^a, \forall a, b \in \mathbb{A} \text{ with } |b| > |a|. \\
\;\text{(iv) } c^a = c^{-a},\ \forall a \in \mathbb{A} \\
\text{and properties (ii), (iii), and (iv) also holding for the } d \text{ and } e \text{ sequences.}
\end{array} \right\} \tag{A.3}$$

be the class of functions that are convex combinations (i) of one-period before history, current and expected future preference shocks, 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, relative to the origin. 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.

Lemma 1 (Convex Combination Form) *For any history (y^0, θ^1), the unique symmetric solution depends solely on last period equilibrium choices and current preference shock realizations, i.e. $y_1^*(y^0, \theta^1) = g_1(y_0, \theta_1)$, for some $g_1 : Y \times \Theta \to Y$. Moreover, the policy function g_1 has the convex combination form as in the statement of the theorem.*

This proves Step 1, namely that the statement of the Theorem is true for 1-period economies. Next, we demonstrate that this result holds for any any finite-horizon, T-period economy.

Step 2: Induction, T-1 implies T. Let $2 \leq T < \infty$. 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. By hypothesis, there exists a unique subgame perfect equilibrium $(g_l^*)_{l=T-1}^1$ for the $T-1$-period continuation economy. Note that we use the notation $l = T-(t-1)$ in the Theorem to denote the time periods remaining until the end of the economy, to make it it easier for the reader. Agent a believes that all other agents, including his own reincarnations, will use that unique symmetric equilibrium map from period 2 on, i.e., for any agent $b \in \mathbb{A}$,

$$y_{b,t}\ \left(y^{t-1}, \theta^t\right) = g_l^*(R^b\, y_{t-1}, R^b\, \theta^t), \qquad \text{for all } l = T-1, \ldots, 1$$

Given any initial history (y^0, θ^1), the current strategies of other agents $(y_{b,1})_{b \neq a}$, and the fact that $(y_{b,t})_{t \geq 2}^{b \in \mathbb{A}}$ are induced by g, agent a solves

$$\begin{aligned} \max_{y_{a,1} \in Y} \quad & \Bigg\{ -\alpha_1 \left(y_{a,0} - y_{a,1}\right)^2 - \alpha_2 \left(\theta_{a,1} - y_{a,1}\right)^2 - \alpha_3 \left(y_{a-1,1} - y_{a,1}\right)^2 - \alpha_3 \left(y_{a+1,1} - y_{a,1}\right)^2 \\ & + E\Bigg[\sum_{\tau=2}^{T} \beta^{\tau-1} \Big(-\alpha_1 \left(y_{a,\tau-1} - y_{a,\tau}\right)^2 - \alpha_2 \left(\theta_{a,\tau} - y_{a,\tau}\right)^2 \\ & -\alpha_3 \left(y_{a-1,\tau} - y_{a,\tau}\right)^2 - \alpha_3 \left(y_{a+1,\tau} - y_{a,\tau}\right)^2\Big) \,\Bigg|\, \left(y^0, \theta^1\right) \Bigg] \Bigg\} \end{aligned} \tag{A.4}$$

Thanks to the linearity of the optimal future choices given by iterative application of equilibrium policy, agent a's problem (A.4) is differentiable with respect to $y_{a,1}$ and the unconstrained ($y_{a,1} \in \mathbb{R}$) first order condition is

$$\begin{aligned} 0 \;=\;& \alpha_1 \left(y_{a,0} - y_{a,1}\right) + \alpha_2 \left(\theta_{a,1} - y_{a,1}\right) + \alpha_3 \left(y_{a-1,1} - y_{a,1}\right) + \alpha_3 \left(y_{a+1,1} - y_{a,1}\right) \\ & + E\Bigg[\sum_{\tau=2}^{T} \beta^{\tau-1} \Big(-\alpha_1\ \left(y_{a,\tau-1} - y_{a,\tau}\right) \frac{\partial}{\partial y_{a,1}} \left(y_{a,\tau-1} - y_{a,\tau}\right) + \alpha_2\ \left(\theta_{a,\tau} - y_{a,\tau}\right) \frac{\partial}{\partial y_{a,1}}\, y_{a,\tau} \\ & -\ \alpha_3\ \left(y_{a-1,\tau} - y_{a,\tau}\right) \frac{\partial}{\partial y_{a,1}} \left(y_{a-1,\tau} - y_{a,\tau}\right) - \alpha_3\ \left(y_{a+1,\tau} - y_{a,\tau}\right) \frac{\partial}{\partial y_{a,1}} \left(y_{a+1,\tau} - y_{a,\tau}\right)\Big) \,\Bigg|\, \left(y^0, \theta^1\right) \Bigg] \end{aligned} \tag{A.5}$$

Moreover, agent a's problem (A.4) is strictly concave in his choice $y_{a,1}$ since the second partial of the objective function with respect to $y_{a,1}$, $-\Delta_T$, is negative

$$\Delta_T \quad := \quad \alpha_1 + \alpha_2 + 2\alpha_3 + \sum_{t=2}^{T} \beta^{t-1} \left(\alpha_1 \left(\frac{\partial}{\partial y_{a,1}} (y_{a,t-1} - y_{a,t}) \right)^2 + \alpha_2 \left(\frac{\partial}{\partial y_{a,1}} y_{a,t} \right)^2 + \alpha_3 \left(\frac{\partial}{\partial y_{a,1}} (y_{a-1,t} - y_{a,t}) \right)^2 + \alpha_3 \left(\frac{\partial}{\partial y_{a,1}} (y_{a+1,t} - y_{a,t}) \right)^2 \right) > 0 \tag{A.6}$$

Consequently, the first order condition (FOC) characterizes the unique maximizer of the unconstrained problem. Since agent's objective (A.4) is a discounted expected weighted sum of quadratic terms with peaks in the set Y, and the conditional expectations of shocks are in the interior of Y, the unique maximizer of the unconstrained problem is almost surely in the interior of Y, as the first part of Lemma 2 states. We prove that lemma in Technical Appendix F.

By hypothesis, there exists a unique subgame perfect equilibrium for the $T-1$-period continuation economy. Hence, iterated application of the policy maps backwards towards period 1 allows one to write any t-period equilibrium choice of agent $a \in \mathbb{A}$ as

$$\begin{aligned}
y_{a,t} &= g_{T-(t-1)}(R^a\, y_{t-1}, R^a\, \theta^t) \\
&= \sum_{b_1 \in \mathbb{A}} c_{b_1,T-(t-1)}\, y_{a+b_1,t-1} + \sum_{b_1 \in \mathbb{A}} d_{b_1,T-(t-1)}\, \theta_{a+b_1,t} + \sum_{b \in \mathbb{A}} \sum_{\tau=t+1}^{T} e_{b,T-(t-1),\tau-t}\, E\left[\theta_{a+b,\tau} | \theta^t\right] \\
&= \sum_{b_1 \in \mathbb{A}} c_{b_1,T-(t-1)}\, \underbrace{g_{T-t}(R^{a+b_1}\, y_{t-2}, R^{a+b_1}\, \theta^{t-1})}_{y_{a+b_1,t-1}} + \sum_{b_1 \in \mathbb{A}} d_{b_1,T-(t-1)}\, \theta_{a+b_1,t} \\
&\quad + \sum_{b \in \mathbb{A}} \sum_{\tau=t+1}^{T} e_{b,T-(t-1),\tau-t}\, E\left[\theta_{a+b,\tau} | \theta^t\right] \\
&\;\vdots \\
&= \sum_{b_1 \in \mathbb{A}} \cdots \sum_{b_{t-1} \in \mathbb{A}} c_{b_1,T-(t-1)} \cdots c_{b_{t-1},T-1}\, y_{a+b_1+\cdots+b_{t-1},1} \\
&\quad + \sum_{s=1}^{t-1} \sum_{b_1 \in \mathbb{A}} \cdots \sum_{b_{s-1} \in \mathbb{A}} c_{b_1,T-(t-1)} \cdots c_{b_{s-1},T-(t-s)+1} \Bigg[\sum_{b_s \in \mathbb{A}} d_{b_s,T-(t-s)} \theta_{a+b_1+\cdots+b_s,t-(s-1)} \\
&\quad + \sum_{b_s \in \mathbb{A}} \sum_{\tau=t-s+2}^{T} e_{b_s,T-(t-s),\tau-(t-s+1)}\, E\left[\theta_{a+b,\tau} | \theta^{t-s+1}\right] \Bigg]
\end{aligned} \tag{A.7}$$

which shows that each future choice can be written as a convex combination of period-1 choices, period-1 shocks, and future expected shocks by iterated application of policy maps. Since at each iteration, convex combination structure is preserved, it is so at the end too. As a direct consequence of (A.7), we have for any $t > 1$ and for $a \in \mathbb{A}$

$$\frac{\partial y_{0,t}}{\partial y_{a,1}} = \sum_{b_1} \cdots \sum_{b_{t-1}} c_{b_1,T-(t-1)} \cdots c_{a-(b_1+\cdots+b_{t-1}),T-1} \tag{A.8}$$

Consequently, the coefficient multiplying $y_{a+b,1}$ in (A.5), $\gamma_{b,T}$, can be obtained by computing the cross partial of the objective function with respect to $y_{a+b,1}$ and $y_{a,1}$ (i.e. the partial of the right hand side of

(A.5)), and it represents the total effect of a change in $y_{a+b,1}$ on the expected discounted marginal utility of agent a. Namely, for any $b \in \mathbb{A}$

$$\begin{aligned} \gamma_{b,T} \quad := \quad & \alpha_3 \, I_{\{b \in \{-1,1\}\}} \\ & - \sum_{t=2}^{T} \beta^{t-1} \left(\alpha_1 \, \frac{\partial}{\partial y_{a+b,1}} \left(y_{a,t-1} - y_{a,t} \right) \frac{\partial}{\partial y_{a,1}} \left(y_{a,t-1} - y_{a,t} \right) + \alpha_2 \, \frac{\partial}{\partial y_{a+b,1}} \, y_{a,t} \, \frac{\partial}{\partial y_{a,1}} \, y_{a,t} \right. \\ & \left. + \, \alpha_3 \, \frac{\partial}{\partial y_{a+b,1}} \left(y_{a-1,t} - y_{a,t} \right) \frac{\partial}{\partial y_{a,1}} \left(y_{a-1,t} - y_{a,t} \right) + \alpha_3 \, \frac{\partial}{\partial y_{a+b,1}} \left(y_{a+1,t} - y_{a,t} \right) \frac{\partial}{\partial y_{a,1}} \left(y_{a+1,t} - y_{a,t} \right) \right). \end{aligned} \tag{A.9}$$

Similarly, the coefficients multiplying $E\left[\theta_{a+b,\tau}|\theta^t\right] =: z(a+b,\tau)$ in equation (A.5)

$$\begin{aligned} \mu_{b,\tau,T} \quad = \quad & \frac{\partial}{\partial z(a+b,\tau)} \, E\left[\sum_{\tau=2}^{T} \beta^{\tau-1} \left(-\alpha_1 \, \left(y_{a,\tau-1} - y_{a,\tau} \right) \frac{\partial}{\partial y_{a,1}} \left(y_{a,\tau-1} - y_{a,\tau} \right) + \alpha_2 \, \left(\theta_{a,\tau} - y_{a,\tau} \right) \frac{\partial}{\partial y_{a,1}} \, y_{a,\tau} \right. \right. \\ & \left. \left. - \, \alpha_3 \, \left(y_{a-1,\tau} - y_{a,\tau} \right) \frac{\partial}{\partial y_{a,1}} \left(y_{a-1,\tau} - y_{a,\tau} \right) - \alpha_3 \, \left(y_{a+1,\tau} - y_{a,\tau} \right) \frac{\partial}{\partial y_a} \left(y_{a+1,\tau} - y_{a,\tau} \right) \right) \middle| \left(y^0, \theta^1 \right) \right] \end{aligned}$$

Moreover, the second part of Lemma 2 states that thanks to the linearity of the FOC in the choice variables, the preference shocks, and the expected future preference shocks, one can write the FOC in equation (A.5) as a function only of contemporaneous choices, and expected future shocks, through iterative application of the policy functions for future period equilibrium choices, as we demonstrated in (A.7). Since the unique optimizer $y_{a,1}$ is almost surely interior, the coefficients multiplying these are non-negative.

Lemma 2 (Interiority) *Let $T \geq 2$. The unique optimizer $y_{a,1}$ is almost surely in the interior of $Y = [\underline{y}, \bar{y}]$, and equation (A.5) can be written as*

$$0 \quad = \quad -y_{a,1} \, \Delta_T + \alpha_1 \, y_{a,0} + \alpha_2 \, \theta_{a,1} + \sum_{b \neq 0} \gamma_{b,T} \, y_{a+b,1} + \sum_{b \in \mathbb{A}} \sum_{\tau=2}^{T} \mu_{b,\tau,T} \, E\left[\theta_{a+b,\tau}|\theta^1\right] \tag{A.10}$$

where $\Delta_T := \alpha_1 + \alpha_2 + \sum_{b \neq 0} \gamma_{b,T} + \sum_{b \in \mathbb{A}} \sum_{\tau=2}^{T} \mu_{b,\tau,T}$, and the coefficients α_1, α_2, $(\gamma_{b,T})_{b \neq 0}$, and $(\mu_{b,\tau,T})_{b \in \mathbb{A}}^{\tau \geq 2}$ are non-negative.

By isolating the choice $y_{a,1}$, we can write the unique maximizer as a convex combination of $y_{a,0}$, $\theta_{a,1}$, $(y_{a+b,1})_{b \neq 0}$ and $(E\left[\theta_{a+b,\tau}|\theta^1\right])_{b \in \mathbb{A}}$

$$y_{a,1} \quad = \quad \Delta_T^{-1} \left(\alpha_1 \, y_{a,0} + \alpha_2 \, \theta_{a,1} + \sum_{b \neq 0} \gamma_{b,T} \, y_{a+b,1} + \sum_{b \in \mathbb{A}} \sum_{\tau=2}^{T} \mu_{b,\tau,T} \, E\left[\theta_{a+b,\tau}|\theta^1\right] \right) \tag{A.11}$$

Each of these are elements of Y, a convex set. Thus, the optimal choice of the unconstrained problem is in the feasible set of the constrained problem, hence it is its unique maximizer. The form in (A.11) 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 $L_T : \mathbf{B} \to \mathbf{B}$ that acts on the family of bounded measurable functions $f = (f_a) \in \mathbf{B} := \prod_{a \in \mathbb{A}} B\left((Y \times \Theta)^s, Y\right)$ according to

$$(L_T \, f)_a \left(y^0, \theta^1\right) = \Delta_T^{-1} \left(\alpha_1 \, y_{a,0} + \alpha_2 \, \theta_{a,1} + \sum_{b \neq 0} \gamma_{b,T} \, f_{a+b} \left(y^0, \theta^1\right) + \sum_{b \in \mathbb{A}} \sum_{\tau=2}^{T} \mu_{b,\tau,T} \, E\left[\theta_{a+b,\tau}|\theta^1\right] \right)$$

Clearly L_T is a self-map. Using straightforward modifications of the arguments in the proof of **Step 1**, one obtains for $f, f' \in \mathbf{B}$ and for all $\left(y^0, \theta^1\right)$ that

$$\left|\left(L_T f\right)_a\left(y^0, \theta^1\right) - \left(L_T f'\right)_a\left(y^0, \theta^1\right)\right| \quad \leq \quad \sum_{b \neq 0}\left(\frac{\gamma_{b,T}}{\Delta_T}\right)\|f - f'\|_\infty$$

The coefficient $\sum_{b\neq 0}\left(\frac{\gamma_{b,T}}{\Delta_T}\right) < 1$ since $\alpha_1 + \alpha_2 > 0$. Thus, L_T is a contraction mapping on the Banach space $\mathbf{B}$; hence it has a unique fixed point f^*. Using straightforward modifications of the arguments in **Step 1**, the unique fixed point f^* is necessarily symmetric. Moreover, L_T maps the subspace (call it B_G) of bounded measurable functions that assume the convex combination form into itself, as we show next. Let $g \in B_G$ be such that after any history $(y^0, \theta^1) = (y_{-(s-1)}, \theta_{-(s-2)}, \ldots, y_{-1}, \theta_0, y_0, \theta_1)$, one has $y_1(y^0, \theta^1) = g(y_0, \theta^1)$ with (c, d, e) being the coefficient sequence associated with g. Applying L_T to g, we get

$$\begin{aligned}\left(L_T\, g\right)\left(R^a y^{t-1}, R^a\theta^t\right) \quad = \quad & \Delta_T^{-1}\Bigg[\underbrace{[\alpha_1 + \sum_{b\neq 0}\gamma_{b,T}\, c_{-b}]}_{\Delta_T c'_0}\, y_{a,0} + \underbrace{[\alpha_2 + \sum_{b\neq 0}\gamma_{b,T}\, d_{-b}]}_{\Delta_T d'_0}\,\theta_{a,1} \\ & + \sum_{b_1 \neq 0}\Bigg(\underbrace{[\sum_{b\neq 0}\gamma_{b,T}\, c_{b_1 - b}]}_{\Delta_T c'_{b_1}}\, y_{a+b_1,0} + \underbrace{[\sum_{b\neq 0}\gamma_{b,T}\, d_{b_1-b}]}_{\Delta_T d'_{b_1}}\,\theta_{a+b,1}\Bigg) \\ & + \sum_{b_1\in\mathbb{A}}\sum_{\tau=2}^{T}\underbrace{[\mu_{b,\tau,T} + \sum_{b\neq 0}\gamma_{b,T}\, e_{b-b_1,\tau-1}]}_{\Delta_T e'_{b_1,\tau-1}}\, E\left[\theta_{a+b_1,\tau}|\theta^1\right]\Bigg] \end{aligned} \tag{A.12}$$

By construction, the policy coefficient series involved in the rearrangement, are non-negative, bounded-valued, absolutely summable sequences. Hence, all sums above converge, since choices and shocks come from the same compact interval of the real line. Consequently, by Fubini's Theorem (see e.g., Dunford and Schwartz (1958), p.190), one may switch the order of summation since the double sums yield a finite answer when the summand is replaced by its absolute value. Moreover, the expression above is linear in period 0 choices, period-1 shocks, and future expected shocks. By definition of the new coefficient sequence (c', d', e'), 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 total sum of the coefficients on the right hand side of (A.12) is $\Delta_T^{-1}(\alpha_1 + \alpha_2 + \sum_{b\neq 0}\gamma_{b,T} + \sum_{b\in\mathbb{A}}\sum_{\tau=2}^{T}\mu_{b,\tau,T}) = 1$ since (A.12) is basically a convex combination of elements and of functions that are convex combinations of elements of the convex set Y. The proof of the properties (i), (ii), (iii), and (iv) follows analogous arguments as in Lemma 1. Thus, the unique fixed point, call it g_T^{**}, lies in the set B_G.

Therefore, when the symmetric continuation equilibrium policies are in G, after any history (y^0, θ^1), the unique symmetric equilibrium policy in the first period, g_T^{**} is in G too. Now, construct the policy function g^* as $g_T^*(y_0, \theta_1) = g_T^{**}(y_0, \theta_1)$ for any initial (y_0, θ_1); and $g_l^*(y_{t-1}, \theta^t) = g_l(y_{t-1}, \theta^t)$, for all $l = T-1, \ldots, 1$ and for all $t \in \{2, \cdots, T\}$. The function g^* is by construction the unique SPE 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 subgame perfect equilibria tends to a stationary symmetric subgame 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 strategy set. For a fixed discount factor $\beta \in (0,1)$, let $L_\beta := \{\beta_T \in [0,1]^\infty \mid \beta_{T,t} = \beta^{t-1},\ \text{for } 1 \le t \le T, \text{and } \beta_{T,t} = 0,\ \text{for } t > T,\ \text{where } T \in \{1,2,\ldots\} \cup \{\infty\}\}$ 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. We can show that

Lemma 3 (Compactness) *L_β and G endowed with the supnorm are compact metric spaces.*

Now, given $g \in G^\infty$, let $y_a(g)$ be agent a's *strategy induced by* g, i.e., $y_a(g)(y^0, \theta^1) = g_t(R^a\, y_0, R^a\, \theta^1)$, for all $a \in \mathbb{A}$ and all (y^0, θ^1). Define the objective function U for agent a in the class of truncated economies as $U : G^\infty \times L_\beta \times G^\infty$ as

$$U(g^0\ ;\ \beta_T, g) := E\left[\sum_{t=1}^{\infty} \beta_{T,t}\, u\left(y_{a,t-1}(g^0), y_{a,t}(g^0), \{y_{a+b,t}(g)\}_{b\in\{-1,1\}}, \theta_{a,t}\right) \Big| (y_0, \theta_1)\right]$$

where u represents the conformity preferences. Pick $\bar{\theta} \in Y$. Let the feasibility correspondence $\Gamma : L_\beta \times G^\infty \to G^\infty$ be defined for $T < \infty$ as $\Gamma(\beta_T, g) = \{g^0 \in G^\infty \mid g_t^0(y, \theta^t) = \bar{\theta} \in \Theta, \forall t > T, \forall (y, \theta^t) \in Y \times \Theta^t\}$, 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 4 (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 equilibrium policy sequence g^{*T}, define $g^{**T} \in G^\infty$ as

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

G^∞ endowed with the product topology is compact since each G endowed with the supnorm is compact from Lemma 3. Since product topology is metrizable, say with metric d,[47] (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 $g^* \in G^\infty$. Let $M : L_\beta \times G^\infty \to G^\infty$ be the correspondence of maximizers of U given the value of the parameters. Lastly, let $\mathcal{E} : L_\beta \to G^\infty$ be the symmetric equilibrium correspondence for the sequence of finite economies. Since g^{*T_n} is a symmetric subgame perfect equilibrium for any T_n, for all $g^{T_n} \in G^\infty$ we

[47] See Footnote 54 for an example of metrization of product topology.

have

$$\begin{aligned}
U(g^*_{T_n}\ ;\ \beta_{T_n}, g^*_{T_n}) &= E\left[\sum_{t=1}^{\infty}\beta_{T_n,t}\, u\left(y_{0,t-1}(g^{*T_n}), y_{0,t}(g^{*T_n}), \{y_{b,t}(g^{*T_n})\}_{b\in\{-1,1\}}, \theta_{0,t}\right)\ \Big|\, (y^0,\theta^1)\right]\\
&= E\left[\sum_{t=1}^{T_n+1}\beta^{t-1}\, u\left(y_{0,t-1}(g^{*T_n}), y_{0,t}(g^{*T_n}), \{y_{b,t}(g^{*T_n})\}_{b\in\{-1,1\}}, \theta_{0,t}\right)\ \Big|\, (y^0,\theta^1)\right]\\
&\geq E\left[\sum_{t=1}^{T_n+1}\beta^{t-1}\, u\left(y_{0,t-1}(g^{T_n}), y_{0,t}(g^{T_n}), \{y_{b,t}(g^{*T_n})\}_{b\in\{-1,1\}}, \theta_{0,t}\right)\ \Big|\, (y^0,\theta^1)\right]\\
&= E\left[\sum_{t=1}^{\infty}\beta_{T_n,t}\, u\left(y_{0,t-1}(g^{T_n}), y_{0,t}(g^{T_n}), \{y_{b,t}(g^{*T_n})\}_{b\in\{-1,1\}}, \theta_{0,t}\right)\ \Big|\, (y^0,\theta^1)\right]\\
&= U(g_{T_n}\ ;\ \beta_{T_n}, g^*_{T_n})
\end{aligned}$$

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 4 and that the feasibility correspondence Γ is continuous, by the Maximum Theorem (see Berge (1963), p. 115), the correspondence of maximizers, M, is upper hemi-continuous. This implies that if $(\beta_{T_n}, g^{*T_n}) \to (\beta_\infty, g^*)$, then $g^* \in M(\beta_\infty, g^*)$ hence g^* is a symmetric SPE of the infinite-horizon economy. This implies immediately that the equilibrium correspondence $\mathcal{E}$ is upper hemi-continuous too.

Uniqueness of finite-horizon symmetric SPEs imply that $\mathcal{E}$ is single-valued hence continuous for $T<\infty$. Define $\mathcal{F}(\beta_T) := \mathcal{E}(\beta_T)$, for $T<\infty$ and let $\mathcal{F}(\beta_\infty) = g^*$. This way, $\mathcal{F}$ is continuous on the space L_β, which is compact under the supnorm by Lemma 3. Consequently, $\mathcal{F}$ is uniformly continuous. This means, for a given $\epsilon>0$, we can pick $\delta>0$ small enough so that $||\beta_T - \beta_{T'}||_\infty < \delta$ implies $d\left(\mathcal{F}(\beta_T), \mathcal{F}(\beta_{T'})\right) < \frac{\epsilon}{2}$. We know from the previous approximation that for $\beta_T \to \beta_\infty$ there is a subsequence $g^{*T_n} \to g^*$. 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, then, any element T_n of the subsequence and any other element, T' such that $T_n, T' \geq T(\delta)$. We have

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

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 $g^{*T_n} \to g^*$ uniformly. This proves that the whole sequence $g^{*T} \to g^*$ 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 stationary SPE policy, hence g^* is stationary. This concludes **Step 3** which in turn establishes the proof of the statement of Theorem 1. ■

B Proof of Theorem 2: Characterization and Computation

-(i) : The first-order conditions of each agent's objective is linear in the choices, the preference shocks, and the conditional expected values of the future shocks. The structural coefficients multiplying these are independent of the realizations or the probability law of the stochastic process θ. Hence, the resulting equilibrium coefficients are so as well.

-(ii) : In the terminal time period, $l = 1$, matching the coefficients of the policy function on both sides of equation (A.2), characterizing the unique optimal choice for agent zero, one obtains for $a \in \mathbb{A}$

$$c_{a,1} = \left(\frac{\alpha_1}{\Delta_1}\right) I_{\{a=0\}} + \left(\frac{\alpha_3}{\Delta_1}\right) c_{a-1,1} + \left(\frac{\alpha_3}{\Delta_1}\right) c_{a+1,1} \tag{B.1}$$

We simply show that one can fit an exponentially declining sequence into this equation. Since the equation has a unique solution as argued in the existence proof, this would prove the statement. If $\alpha_1\alpha_3 = 0$, equation (B.1) implies that $c_{a,1} = 0$ for all $a \neq 0$ and $c_{0,1} = \alpha_1/(\alpha_1 + \alpha_2)$. So, the statement is trivially satisfied where the rate of decline is zero. Assume now that $\alpha_1\alpha_3 \neq 0$. Assume wlog that $a > 0$, we can safely divide both sides by $c_{a-1,1}$ now and multiply them by $\left(\frac{\Delta_1}{\alpha_3}\right)$ to obtain

$$\left(\frac{\Delta_1}{\alpha_3}\right) \underbrace{\left(\frac{c_{a,1}}{c_{a-1,1}}\right)}_{r_1} = 1 + \underbrace{\left(\frac{c_{a+1,1}}{c_{a,1}}\right)}_{r_1} \underbrace{\left(\frac{c_{a,1}}{c_{a-1,1}}\right)}_{r_1}$$

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 $\alpha_1 + \alpha_2 > 0$). The equation has two positive roots, one bigger and one smaller than 1. The bigger root cannot work since it is explosive as $|a| \to \infty$. We pick the smaller root

$$0 < r_1 = \left(\frac{\Delta_1}{2\,\alpha_3}\right) - \sqrt{\left(\frac{\Delta_1}{2\,\alpha_3}\right)^2 - 1} < 1 \tag{B.2}$$

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

$$\sum_{a \in \mathbb{A}} c_{a,1} = \sum_{a \in \mathbb{A}} c_{0,1}\, r_1^{|a|} = c_{0,1} + 2\, c_{0,1}\, \frac{r_1}{1 - r_1} = \frac{\alpha_1}{\alpha_1 + \alpha_2} \tag{B.3}$$

Solving for $c_{0,1}$ from above, we obtain

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

and finally thanks to exponentiality

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

The argument for the sequence $(d_{a,1})_{a \in \mathbb{A}}$ is identical with one change: The sum of coefficients $\sum_a d_{a,1} = \left(\frac{\alpha_2}{\alpha_1 + \alpha_2}\right)$. Note that $e_{a,1} = 0$ for $a \in \mathbb{A}$ optimally since there are no future shocks. This proves part (ii) of the theorem.

-(iii) : We will use the following lemma which we prove in Technical Appendix G.

Lemma 5 (Monotone Increasing Cross-Sectional Rates) *For any $l \geq 2$, the rates at which the policy coefficients converge to zero at the cross-section are strictly monotonic in a, i.e., for any $a \in \mathbb{A}$*

$$r_{|a|+1,l} = \frac{c_{|a|+1,l}}{c_{|a|,l}} < \frac{c_{|a|+2,l}}{c_{|a|+1,l}} = r_{|a|+2,l} \tag{B.4}$$

Moreover, given β and $\frac{\alpha_1}{\alpha_1+\alpha_2}$, the cross-sectional rates are strictly increasing in α_3, i.e.,

$$r_{a,l}(\alpha_3') > r_{a,l}(\alpha_3), \qquad \text{for any } a \neq 0. \tag{B.5}$$

The analogous results hold for for d_l and e_l.

Exponential convergence (as in $T = 1$) implies that $r_a = r_b$ for any $a, b \neq 0$. As the Lemma demonstrates, for $T \geq 2$, as $|a|$ increases (going away from zero at the cross-section), coefficients decline much faster for closer agents than for farther agents. In case N is large, the rate of decline stabilizes eventually to a slower exponential rate, $\lim_{|a|\to\infty} r_{a,T} > r_{b,T}$, for any $b \in \mathbb{A}$, due to the fact that $(r_{a,T})_{a\in\mathbb{A}}$ is a bounded monotone increasing sequence in the interval $(0, 1)$.

-(iv) : As defined in section 2.1, $C_l := \sum_b c_{b,l}$ denote the total effect of past actions; that is, the effect on an agent a's action of a uniform unitary increase in all agents' past actions. Similarly, $D_l := \sum_b d_{b,l}$ denote the total effect of contemporary preference shocks; and $E_l := \sum_b e_{b,l}$ the total effect of expected future preference shocks. We demonstrate in the next Lemma, wich we prove in the Technical Appendix G, that these sums are given by a continued fraction form across periods.

Lemma 6 (Policy Coefficient Sums) *For a T-period dynamic conformity economy with $T > 1$, the policy coefficient sums for $l = 2, \ldots, T$ are given by the following recursive system of continued fractions*

$$\begin{aligned} C_l &= \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1\beta\,(1 - C_{l-1})} \\ D_l &= \frac{\alpha_2}{\alpha_1 + \alpha_2 + \alpha_1\beta\,(1 - C_{l-1})} \\ E_l &= \frac{\alpha_1\beta\,(1 - C_{l-1})}{\alpha_1 + \alpha_2 + \alpha_1\beta\,(1 - C_{l-1})} \end{aligned} \tag{B.6}$$

where $C_1 = \frac{\alpha_1}{\alpha_1+\alpha_2}$, $D_1 = \frac{\alpha_2}{\alpha_1+\alpha_2}$, and $E_1 = 0$. Moreover, $C_l \downarrow C_\infty$ and $D_l \downarrow D_\infty$ are monotonically decreasing (hence $E_l \uparrow E_\infty$) sequences where C_∞, D_∞, and E_∞ are the fixed points of the respective equations in the recursive system (B.6).

Using the structure in Lemma 6, for the period just before the last ($l = 2$),

$$C_2 = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1\beta(1 - C_1)} \leq \frac{\alpha_1}{\alpha_1 + \alpha_2} = C_1$$

and strictly so when $\alpha_1\beta > 0$. Assuming now that $C_{l-1} < \ldots < C_1$, we also obtain

$$C_l = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1\beta(1 - C_{l-1})} < \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1\beta(1 - C_{l-2})} = C_{l-1}$$

as claimed in the statement. The proof for D_l is identical. Since, $E_l = 1 - C_l - D_l$ and that C_l and D_l increase as l decreases, the total effect of expected future preference shocks, E_l instead decreases as l decreases, as stated.

C Proof of Theorem 3: Dependence on the Preference Parameters

- *Proof of* (i) : We prove here injectivity of the maps going from structural parameters into the space of policy coefficients. We start with the finite horizon economy.

- *Finite Horizon*: $1 < T < \infty$. We will first demonstate that the map $(\alpha_1, \alpha_3) \to (c_{b,1}(\alpha_1, \alpha_3))_b$ is injective. Since $\alpha_1 + \alpha_2 > 0$ by Assumption 1, Lemma 6 (see also the proof of Theorem 2-(ii)) yields $C_1 = \frac{\alpha_1}{\alpha_1+\alpha_2}$, and $D_1 = \frac{\alpha_2}{\alpha_1+\alpha_2}$; hence, $C_1 + D_1 = \sum_a c_{a,1} + \sum_a d_{a,1} = \frac{\alpha_1}{\alpha_1+\alpha_2} + \frac{\alpha_2}{\alpha_1+\alpha_2} = 1$. Pick the coefficients whose sum is non-zero. Since the structural equations are the same, the arguments below are the same. So, assume wlog that $C_1 > 0$. Now, there are two possibilities:

(i) If $c_{b,1} = 0$ for all $b \neq 0$, Theorem 2-(ii) (see also footnote 19) implies that $\alpha_3 = 0$. In that case, we can also recover $\alpha_1 = C_1$ and $\alpha_2 = C_2$.

(ii) If there is $b \neq 0$ such that $c_{b,1} \neq 0$, then what we know from Theorem 2-(ii) (see also footnote 19), namely

$$c_{b,1} = \left(\frac{\alpha_1}{1-2\alpha_3}\right)\left(\frac{1-r_1}{1+r_1}\right) r_1^{|b|} > 0, \qquad \text{for all } b \in \mathbb{A}$$

$$\text{where} \qquad r_1 = \left(\frac{1}{2\,\alpha_3}\right) - \sqrt{\left(\frac{1}{2\,\alpha_3}\right)^2 - 1} < 1$$

would imply that $c_{b,1} \neq 0$ for all $b \in \mathbb{A}$. Moreover, we can recover r_1 by computing the ratio of two consecutive policy coefficients, for instance, $r_1 = c_{1,1}/c_{0,1}$. Since, r_1, as shown in the second equation above, is strictly increasing in α_3, there exists a unique value of α_3 that generates the policy coefficient sequence we observe. Moreover, knowing the true value of α_3, we can also recover $\alpha_1 = (1-2\alpha_3)\,C_1$ and $\alpha_2 = (1-2\alpha_3)(1-C_1)$.

So far, we established that the map $(\alpha_1, \alpha_3) \to (c_{b,1}(\alpha_1, \alpha_3))_b$ is injective. Now, let an observationally equivalent series of policy coefficients $(c_{b,l}(\alpha_1, \alpha_3, \beta))_{b\in\mathbb{A}, l>1}$ and $(c_{b,l}(\alpha'_1, \alpha'_3, \beta'))_{b\in\mathbb{A}, l>1}$ be given. Namely, for all $l > 1$ and any $b \in \mathbb{A}$

$$c_{b,l}(\alpha_1, \alpha_3, \beta) = c_{b,l}(\alpha'_1, \alpha'_3, \beta')$$

This implies that the sums should match for any period as well, i.e., for any $l > 1$

$$C_l = \sum_{b\in\mathbb{A}} c_{b,l}(\alpha_1, \alpha_3, \beta) = \sum_{b\in\mathbb{A}} c_{b,l}(\alpha'_1, \alpha'_3, \beta') = C'_l$$

Let $p := \frac{\alpha_1}{\alpha_1+\alpha_2}$ and $p' := \frac{\alpha'_1}{\alpha'_1+\alpha'_2}$. So, we can write $C_1 = p$ and $C'_1 = p'$. Moreover, by dividing the numerator and the denominator of the right hand side of the continued fractions in Lemma 6 by $(\alpha_1+\alpha_2)$, we can write for $l > 1$

$$C_l = \frac{p}{1+\beta\, p\,(1-C_{l-1})} = \frac{p'}{1+\beta'\, p'\,(1-C'_{l-1})} = C'_l \tag{C.1}$$

Given the observable sequence of sums $(C_l)_{l>1}$, one can solve for the unique value of p for any given value of β from (C.1) as

$$p \quad = \quad \frac{C_l}{1 \,-\, \beta\, C_l\,(1-C_{l-1})}$$

This establishes a continuous, strictly monotonically increasing, and differentiable function between β and p and identifies the following sets of observationally equivalent (p,β) pairs for any $l>1$

$$\mathcal{P}_l \quad := \quad \{(p'',\beta'') \in [0,1]\times[0,1) \;\text{ such that }\; C_l(p'',\beta'') = C_l\}$$

Now, we will show that equilibrium dynamics across periods impose sufficient restrictions on the inverse image of the observed policy coefficient sequences. A total derivative, assuming that the levels C_l and C_{l-1} do not change, yields

$$\frac{\partial p}{\partial \beta}\Big|_{\mathcal{P}_l} \quad = \quad p^2\,(1-C_{l-1})$$

We know from Theorem 2-(iv) that for $l>1$

$$p = C_1 > \ldots > C_l > \ldots$$

Hence, for any $l' > l > 1$

$$\frac{\partial p}{\partial \beta}\Big|_{\mathcal{P}_{l'}} \;=\; p^2\,(1-C_{l'-1}) \quad > \quad p^2\,(1-C_{l-1}) \;=\; \frac{\partial p}{\partial \beta}\Big|_{\mathcal{P}_l} \tag{C.2}$$

Since the parameter vectors $(\alpha_1,\alpha_3,\beta)$ and $(\alpha'_1,\alpha'_3,\beta')$ generate these sequence of sums, $(p,\beta)\in\mathcal{P}_l$ and $(p',\beta')\in\mathcal{P}_l$, for any $l>1$. As seen in (C.2), the marginal rates of substitution between p and β, at a given point in $\cap_{l>1}\mathcal{P}_l$, are ranked across the sets $\mathcal{P}_l$ and $\mathcal{P}_{l'}$. So, these two sets can cross only once. Hence, knowing the value of C_l and $C_{l'}$ for any $l'>l>1$ pins down the unique pair, say (p^*,β^*) that can generate the observable coefficients. This implies that

$$p \;=\; p' \;=\; p^* \qquad \text{and} \qquad \beta \;=\; \beta' \;=\; \beta^*$$

We know from Lemma 5 (also used in the proof of Theorem 2-(iii)) that cross-sectional rates of convergence of the policy coefficients are strictly increasing in α_3, given β and $p = \frac{\alpha_1}{\alpha_1+\alpha_2}$. Hence, we can read $\left(\frac{c_{|a|+1,l}}{c_{|a|,l}}\right)$ by computing the ratio of two consecutive policy coefficients, at the cross-section. Thanks to Lemma 5, there exists a unique value of α_3 that generates that ratio value. Given that we can determine the value of α_3, we can also recover $\alpha_1 = p\,(1-2\alpha_3)$ and $\alpha_2 = (1-p)\,(1-2\alpha_3)$. This establishes that the map $(\alpha_1,\alpha_3,\beta) \to (c_{b,l}(\alpha_1,\alpha_3,\beta))_{b\in\mathbb{A}}^{l>1}$ is injective.

- *Infinite Horizon*: Let $T=\infty$. Our objective is to demonstate that the map $(\alpha_1,\alpha_3,\beta)\to(c_{b,1}(\alpha_1,\alpha_3,\beta))_b$ is injective. It will be easier to use a change of variable, as in the finite horizon case. Namely, given that $p := \frac{\alpha_1}{\alpha_1+\alpha_2}$, we can write $\alpha_2 = (1-p)(\alpha_1+\alpha_2) = \alpha_1\left(\frac{1-p}{p}\right)$. We also use the normalization $\alpha_1+\alpha_2+2\alpha_3 = 1$ to write $\alpha_3 = \frac{1}{2}(1-(\alpha_1+\alpha_2)) = \left(\frac{p-\alpha_1}{2p}\right)$. With this change of variable, our paramater vector of interest becomes (α_1,p,β).

We cannot use the identification power of the equilibrium dynamics across periods since we have access only to the stationary policy coefficients. Therefore, instead of looking across periods, we need to look for variation across agents. To do that, we study more closely the equation system, derived from the first order conditions of any agent $a \in \mathbb{A}$, that generates the policy coefficients. Thanks to Lemma 2 and by matching equilibrium coefficients c in (A.10), one obtains for any $b \in \mathbb{A}$

$$0 \;=\; -\gamma_0\, c_0 \;+\; \alpha_1\, I_{\{b=0\}} \;+\; \sum_{b_1 \neq 0} \gamma_{b_1}\, c_{b-b_1} \tag{C.3}$$

where $\gamma_0 := \Delta_\infty$ and γ_{b_1} is as we defined in (A.9): the total effect on agent a's expected discounted marginal utility, of a change in the first period choice, $y_{a+b_1,1}$ of an agent b_1 distance away from agent a, namely agent $a + b_1$.

For the rest of the proof, we will need the following lemma which we prove in Technical Appendix G.

Lemma 7 (Cross-sectional Variation) *There exists a unique sequence* $(\bar{\gamma}_b) := (\alpha_1^{-1}\, \gamma_{b_1})$ *that satisfies the system in (C.3). Moreover,* (i) *the map* $\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1}$ *is continuously differentiable and the partial derivatives satisfy* $\frac{\partial}{\partial p}\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1} > 0$ *and* $\frac{\partial}{\partial \alpha_1}\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1} < 0$*, for any* $(\alpha_1, p, \beta(p))$*; and* (ii) *there exist agents* $b \neq b'$ *such that for any* $(\alpha_1, p, \beta(p))$*,*

$$-\frac{\frac{\partial \alpha_1^{-1}\gamma_{b'}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b'}}{\partial \alpha_1}}\Bigg|_{\alpha_1^{-1}\gamma_{b'}(\alpha_1,p,\beta(p))=\bar{\gamma}'_b} \quad > \quad -\frac{\frac{\partial \alpha_1^{-1}\gamma_{b}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b}}{\partial \alpha_1}}\Bigg|_{\alpha_1^{-1}\gamma_{b}(\alpha_1,p,\beta(p))=\bar{\gamma}_b} \quad > \; 0. \tag{C.4}$$

Let an observationally equivalent series of policy coefficients $(c_b(\alpha_1, p, \beta))_{b\in\mathbb{A}}$ and $(c_b(\alpha'_1, p', \beta'))_{b\in\mathbb{A}}$ be given. Namely, for any $b \in \mathbb{A}$

$$c_b(\alpha_1, p, \beta) \quad = \quad c_b(\alpha'_1, p', \beta')$$

and thanks to Lemma 6 the sums should satisfy

$$C_\infty = \sum_{b\in\mathbb{A}} c_b(\alpha_1, p, \beta) \;=\; \frac{p}{1 + \beta\, p\,(1 - C_\infty)} \quad = \quad \frac{p'}{1 + \beta'\, p'\,(1 - C'_\infty)} \;=\; \sum_{b\in\mathbb{A}} c_b(\alpha'_1, p', \beta') = C'_\infty.$$

One can solve for the unique value of β given any value of p for any given level C_∞.

$$\beta(p \,|\, C_\infty) \quad = \quad \frac{p - C_\infty}{p\, C_\infty (1 - C_\infty)} \tag{C.5}$$

which establishes a continuous, strictly monotonically increasing, and differentiable function $\beta(p \,|\, C_\infty)$, and identifies a set of observationally equivalent (p, β) pairs consistent with the observed levels of policy coefficient sum C_∞.

Given that we observe (c_b), the system in (C.3) need to be solved in terms of (γ_{b_1}) and α_1. Stacking the coefficients multiplying unknowns in each equation in (C.3) in a separate row vector forms a circulant matrix.[48] With this definition, as we state in Lemma 7, we can invert that matrix and obtain the unique sequence $(\bar{\gamma}_{b_1}) := (\alpha_1^{-1}\, \gamma_{b_1})$ that solves the system in (C.3). We cannot identify α_1 separately since we have just as many equations as the number of agents in (C.3). This unique solution sequence provides us with $|\mathbb{A}|$ additional parameter restrictions, in addition to the one we obtained in (C.5).

[48]A **circulant matrix** is a special kind of Toeplitz matrix fully specified by one vector, which appears as one of the rows of the matrix. Each other row vector of the matrix is shifted one element to the right relative to the preceding row vector. See e.g. Davis (1970) for an in-depth discussion of circulant matrices.

Next, we replace β by the expression in (C.5). Given that we know the structural form of the mappings $\gamma_{b_1}(\alpha_1, p, \beta)$ from the definition in (A.9), these restrictions jointly identify the following sets of observationally equivalent $(\alpha_1, p, \beta(p))$ vectors for any $b_1 \in \mathbb{A}$,

$$\mathcal{P}_{b_1} \quad := \quad \left\{(\alpha_1'', p'', \beta''(p'')) \in [0,1] \times [0,1] \times [0,1) \text{ such that } \alpha_1^{-1}\gamma_{b_1}(\alpha_1'', p'', \beta''(p'')) = \bar{\gamma}_{b_1}\right\} \quad \text{(C.6)}$$

Thus, we have a total of $|\mathbb{A}|$ restrictions and two parameters (α_1, p) to determine.

Thanks to Lemma 7(i), the Implicit Function Theorem (see e.g., Bartle (1976), p.384) states that, for any $b_1 \in \mathbb{A}$, there exists a continuous and differentiable function $\alpha_1(p \,|\, \mathcal{P}_{b_1})$ that gives the unique value of α_1 for any value of p such that $(\alpha_1, p, \beta(p)) \in \mathcal{P}_{b_1}$, the level set defined in (C.6). The slope of $\alpha_1(p \,|\, \mathcal{P}_{b_1})$ is the marginal rate of substitution between α_1 and p that sustain the same level for the function $\alpha_1^{-1}\gamma_{b_1}(\alpha_1, p, \beta(p))$. Using once again the Implicit Function Theorem, we know that this slope is computed as

$$\frac{d\alpha_1}{dp}\Big|_{\mathcal{P}_{b_1}} \quad = \quad -\frac{\frac{\partial \alpha_1^{-1}\gamma_{b_1}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b_1}}{\partial \alpha_1}}\Bigg|_{\mathcal{P}_{b_1}} \quad > \quad 0$$

and is positive thanks to Lemma 7(i). Therefore, we also know that the implicit functions $\alpha_1(p \,|\, \mathcal{P}_{b_1})$ are strictly increasing in p, for any $b_1 \in \mathbb{A}$. Finally, we know from Lemma 7(ii) that there exist agents $b \neq b'$ such that for any $(\alpha_1, p, \beta(p))$,

$$\frac{d\alpha_1}{dp}\Big|_{\mathcal{P}_{b'}} \quad = \quad -\frac{\frac{\partial \alpha_1^{-1}\gamma_{b'}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b'}}{\partial \alpha_1}}\Bigg|_{\alpha_1^{-1}\gamma_{b'}(\alpha_1,p,\beta(p))=\bar{\gamma}_b'} \quad > \quad -\frac{\frac{\partial \alpha_1^{-1}\gamma_{b}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b}}{\partial \alpha_1}}\Bigg|_{\alpha_1^{-1}\gamma_{b}(\alpha_1,p,\beta(p))=\bar{\gamma}_b} \quad = \quad \frac{d\alpha_1}{dp}\Big|_{\mathcal{P}_{b}}. \quad \text{(C.7)}$$

So, $\alpha_1(p \,|\, \mathcal{P}_{b'})$ is steeper than $\alpha_1(p \,|\, \mathcal{P}_b)$. Hence, $\alpha_1(p \,|\, \mathcal{P}_{b'})$ and $\alpha_1(p \,|\, \mathcal{P}_b)$ can intersect at most once. Moreover, since the parameter vectors (α_1, p, β) and (α_1', p', β') generate the observable sequences of coefficients, $(\alpha_1, p, \beta(p)) \in \mathcal{P}_b \cap \mathcal{P}_{b'}$ and $(\alpha_1', p', \beta(p')) \in \mathcal{P}_b \cap \mathcal{P}_{b'}$. Since $|\mathcal{P}_b \cap \mathcal{P}_{b'}| = 1$, there can only be a unique pair (α_1^*, p^*) that is consistent with the above restrictions. Hence,

$$p \;=\; p' \;=\; p^*, \qquad \text{and } \alpha_1 \;=\; \alpha_1' \;=\; \alpha_1^*$$

This in turn yields the unique value of $\beta^* = \beta(p^* \,|\, C_\infty)$ consistent with the observable level C_∞. Moreover, we can also recover $\alpha_2^* = \left(\frac{1-p}{p}\right)\alpha_1^*$ and $\alpha_3^* = \frac{1}{2}(1 - (\alpha_1^* + \alpha_2^*))$. This establishes that the map $(\alpha_1, \alpha_3, \beta) \to (c_b(\alpha_1, \alpha_3, \beta))_{b \in \mathbb{A}}$ is injective.

- *Proof of* (ii) : As before, we give the arguments for the policy coefficients on history. Same arguments apply to other coefficients as well. Let $(\alpha_1, \alpha_3, \beta)$ and $(\alpha_1', \alpha_3', \beta')$ be given. Assume as in the statement of the theorem that $p = \frac{\alpha_1}{\alpha_1+\alpha_2} = \frac{\alpha_1'}{\alpha_1'+\alpha_2'} = p'$ and $\beta = \beta'$, but $\alpha_3' > \alpha_3$. We know from Lemma 6 that $C_1 = \frac{\alpha_1}{\alpha_1+\alpha_2} = p = p' = \frac{\alpha_1}{\alpha_1+\alpha_2}$, which are independent of the level of α_3 and α_3'. Now assume, for induction, that the sums are equal and independent of α_3 and α_3' up to $l = T - 1 > 0$. By dividing the numerator and the denominator of the right hand side of the continued fractions in Lemma 6 by $(\alpha_1 + \alpha_2)$ and $(\alpha_1' + \alpha_2')$ respectively, we can write

$$C_l \;=\; \frac{p}{1 + \beta\, p\,(1 - C_{l-1})} \quad = \quad \frac{p'}{1 + \beta'\, p'\,(1 - C_{l-1}')} \;=\; C_l'$$

This equation implies that since $C_{l-1} = C'_{l-1}$ are independent of α_3 and α'_3, so are C_l and C'_l. Same argument holds at the limit C_∞ as well. Next, we show the mean preserving spread property. Pick any $l > 0$. We use the symmetry of the policy coefficient sequences around zero to write

$$\begin{aligned} C_l(\alpha'_3) = \sum_{b\in\mathbb{A}} c_{b,l}(\alpha'_3) &= c_{0,l}(\alpha'_3) + 2\sum_{a\geq 1} c_{0,l}(\alpha'_3)\prod_{s=1}^{a} r_{s,l}(\alpha'_3) \\ &= c_{0,l}(\alpha'_3)\left[1 + 2\sum_{a\geq 1}\prod_{s=1}^{a} r_{s,l}(\alpha'_3)\right] \\ &= c_{0,l}(\alpha_3)\left[1 + 2\sum_{a\geq 1}\prod_{s=1}^{a} r_{s,l}(\alpha_3)\right] = \sum_{b\in\mathbb{A}} c_{b,l}(\alpha_3) = C_l(\alpha_3) \end{aligned}$$

Lemma 5 states that $r_{a,l}(\alpha'_3) > r_{a,l}(\alpha_3)$ for all $a \neq 0$. This implies that $\sum_{a\geq 1}\prod_{s=1}^{a} r_{s,l}(\alpha'_3) > \sum_{a\geq 1}\prod_{s=1}^{a} r_{s,l}(\alpha_3)$ in the above equation system. Since the policy coefficient sums are equal, then it has to be the case that

$$\begin{aligned} c_{0,l}(\alpha'_3) &= C_l(\alpha'_3)\left[1 + 2\sum_{a\geq 1}\prod_{s=1}^{a} r_{s,l}(\alpha'_3)\right]^{-1} \\ &< C_l(\alpha_3)\left[1 + 2\sum_{a\geq 1}\prod_{s=1}^{a} r_{s,l}(\alpha_3)\right]^{-1} \\ &= c_{0,l}(\alpha_3) \end{aligned}$$

This shows that the mass at the center decreases as α_3 increases. Now, we will show that the tails of the distribution of coefficients lift up when α_3 increases. We show the argument on the right hand side of zero. Thanks to symmetry, it will hold on both sides. Let us first define $\bar{a} := \inf_{a>0}\{c_{a,l}(\alpha'_3) > c_{a,l}(\alpha_3)\}$. This set should be nonempty. Otherwise, it would mean that $c_{a,l}(\alpha'_3) \leq c_{a,l}(\alpha_3)$ for all a and with strict inequality for $a = 0$, as we showed above. This in turn would mean that $C_l(\alpha'_3) = \sum_a c_{a,l}(\alpha'_3) < \sum_a c_{a,l}(\alpha_3) = C_l(\alpha_3)$, which would be a contradiction. So, the claim is true and $\bar{a}$ exists. But then, by definition of $\bar{a}$, the following holds:

$$\begin{aligned} c_{\bar{a},l}(\alpha'_3) &= c_{0,l}(\alpha'_3)\prod_{s=1}^{\bar{a}} r_{s,l}(\alpha'_3) \\ &> c_{0,l}(\alpha_3)\prod_{s=1}^{\bar{a}} r_{s,l}(\alpha_3) \end{aligned}$$

Hence, for any $a > \bar{a}$,

$$\begin{aligned} c_{a,l}(\alpha'_3) &= c_{0,l}(\alpha'_3)\prod_{s=1}^{\bar{a}} r_{s,l}(\alpha'_3)\prod_{s_1=\bar{a}+1}^{a} r_{s_1,l}(\alpha'_3) \\ &> c_{0,l}(\alpha_3)\prod_{s=1}^{\bar{a}} r_{s,l}(\alpha_3)\prod_{s_1=\bar{a}+1}^{a} r_{s_1,l}(\alpha_3) \\ &= c_{a,l}(\alpha_3) \end{aligned}$$

The first line is by definition; the second is by the monotonicity of the rates in α_3. So, we showed that for all $0 \leq a < \bar{a}$, $c_{a,l}(\alpha'_3) \leq c_{a,l}(\alpha_3)$ with strict inequality for $a = 0$; and for all $a \geq \bar{a}$, $c_{a,l}(\alpha'_3) > c_{a,l}(\alpha_3)$.

Thanks to symmetry around zero, this shows that seeing the policy coefficients as probability distributions on $\mathbb{A}$, $\left(\frac{c_{a,l}(\alpha_3')}{C_l(\alpha_3')}\right)_{a\in\mathbb{A}}$ is a mean-preserving spread of $\left(\frac{c_{a,l}(\alpha_3)}{C_l(\alpha_3)}\right)_{a\in\mathbb{A}}$. This concludes the proof.

D Proof of Theorem 4: Ergodicity

Suppose that the process $\left((\theta_t^a)_{t=-\infty}^{\infty}\right)_{a\in\mathbb{A}}$ is (s)-Markovian. Let π be the initial measure on the configuration space $\mathbf{Y}$ which is the distribution of

$$\begin{aligned} y_0 &= \left(\sum_{m=1}^{\infty}\sum_{b_1}\cdots\sum_{b_m} c_{b_1}\cdots c_{b_{m-1}}\Big[d_{b_m}\,\theta_{a+b_1+\cdots+b_m,1-m}\right. \\ &+ \left.\sum_{\tau=1}^{\infty} e_{b_m,\tau}\,E[\theta_{a+b_1+\ldots+b_m,1-m+\tau}\,|\,\theta_{1-m},\ldots,\theta_{1-m-s}]\Big]\right)_{a\in\mathbb{A}} \end{aligned} \tag{D.1}$$

$(y_t\in\mathbf{Y})_{t=0}^{\infty}$ is an equilibrium process generated by the symmetric Subgame Perfect equilibrium policy function g^*. Hence, given y_0, one obtains on the equilibrium path

$$\begin{aligned} y_{a,1} &= \sum_{b_1} c_{b_1}\,y_{a+b_1,0} + \sum_{b_1} d_{b_1}\,\theta_{a+b_1,1} + \sum_{b_1}\sum_{\tau=1}^{\infty} e_{b_1}\,E[\theta_{a+b_1,1+\tau}\,|\,\theta_1,\ldots,\theta_{1-s}] \\ &= \sum_{b_1} c_{b_1}\left(\sum_{m=1}^{\infty}\sum_{b_1}\cdots\sum_{b_m} c_{b_1}\cdots c_{b_{m-1}}\Big[d_{b_m}\,\theta_{a+b_1+\cdots+b_m,1-m}\right. \\ &\quad + \left.\sum_{\tau=1}^{\infty} e_{b_m,\tau}\,E[\theta_{a+b_1+\ldots+b_m,1-m+\tau}\,|\,\theta_{1-m},\ldots,\theta_{1-m-s}]\Big]\right) \\ &\quad + \sum_{b_1} d_{b_1}\,\theta_{a+b_1,1} + \sum_{b_1}\sum_{\tau=1}^{\infty} e_{b_1}\,E[\theta_{a+b_1,1+\tau}\,|\,\theta_t,\ldots,\theta_{t-s}] \\ &= \sum_{m=1}^{\infty}\sum_{b_1}\cdots\sum_{b_m} c_{b_1}\cdots c_{b_{m-1}}\Big[d_{b_m}\,\theta_{a+b_1+\cdots+b_m,2-m} \\ &\quad + \sum_{\tau=1}^{\infty} e_{b_m,\tau}\,E[\theta_{a+b_1+\ldots+b_m,1-m+\tau}\,|\,\theta_{2-m},\ldots,\theta_{2-m-s}]\Big] \end{aligned}$$

which has the same form as in (D.1). Since the process $\left((\theta_t^a)_{t=-\infty}^{\infty}\right)_{a\in\mathbb{A}}$ is (s)-Markovian, $y_{a,0}$ and $y_{a,1}$ 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 $(y_t)_{t=0}^{\infty}$. Moreover, iterative application of the stationary policy function $g*$ on any path $(\theta_1,\theta_2,\ldots)$ of the stochastic process yields

$$\begin{aligned} y_{a,t} &= \sum_{b_1} c_{b_1}\,y_{a+b_1,t-1} + \sum_{b_1} d_{b_1}\,\theta_{a+b_1,t} + \sum_{b_1}\sum_{\tau=1}^{\infty} e_{b_1}\,E[\theta_{a+b_1,t+\tau}\,|\,\theta_t,\ldots,\theta_{t-s}] \\ &\vdots \\ &= C^t\sum_{b_1}\cdots\sum_{b_t}\left(\frac{c_{b_1}\cdots c_{b_t}}{C^t}\right)y_{a+b_1+\cdots+b_t,0} + \sum_{m=1}^{t}\sum_{b_1}\cdots\sum_{b_m} c_{b_1}\ldots c_{b_{m-1}}d_{b_m}\,\theta_{a+b_1+\ldots+b_m,t+1-m} \\ &\quad + \sum_{m=1}^{t}\sum_{b_1}\cdots\sum_{b_m}\sum_{\tau=1}^{\infty} c_{b_1}\ldots c_{b_{m-1}}e_{b_m,\tau}\,E[\theta_{a+b_1+\ldots+b_m,t+1-m+\tau}\,|\,\theta_{t+1-m},\ldots,\theta_{t-s+1-m}] \end{aligned}$$

Since the preference shock process is stationary (s)-Markov, the law for the sum of preference shocks and expectations is identical to the law for its 't-translated-into-the-past' version, i.e., that of

$$\begin{aligned} y_{a,t} &= C^t \sum_{b_1} \cdots \sum_{b_t} \left(\frac{c_{b_1} \cdots c_{b_t}}{C^t} \right) y_{a+b_1+\cdots+b_t,0} \\ &+ \sum_{m=1}^{t} \sum_{b_1} \cdots \sum_{b_m} c_{b_1} \ldots c_{b_{m-1}} d_{b_m} \, \theta_{a+b_1+\ldots+b_m,1-m} \\ &+ \sum_{m=1}^{t} \sum_{b_1} \cdots \sum_{b_m} \sum_{\tau=1}^{\infty} c_{b_1} \ldots c_{b_{m-1}} e_{b_m,\tau} \, E[\theta_{a+b_1+\ldots+b_m,1-m+\tau} \,|\, \theta_{1-m}, \ldots, \theta_{1-m-s}] \end{aligned} \tag{D.2}$$

$C^t \to 0$ as $t \to \infty$ since $C < 1$ due to the fact that $\alpha_1 + \alpha_2 > 0$. 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 $t \to \infty$. Since the equilibrium is symmetric, the convergence is uniform across agents: $y_t \to y = (y_a)$ uniformly. Thus, for any given initial value y_0, and a path $(\ldots, \theta_{-1}, \theta_0)$, the pointwise limit of $y_{a,t}$ can be written as

$$\begin{aligned} y_a &= \sum_{m=1}^{\infty} \sum_{b_1} \cdots \sum_{b_m} c_{b_1} \cdots c_{b_{m-1}} \Big[d_{b_s} \, \theta_{a+b_1+\cdots+b_s,1-m} \\ &\quad + \sum_{\tau=1}^{\infty} e_{b_m,\tau} \, E[\theta_{a+b_1+\ldots+b_m,1-m+\tau} \,|\, \theta_{1-m}, \ldots, \theta_{1-m-s}] \Big] \end{aligned} \tag{D.3}$$

Now, pick any $f \in C(\mathbf{Y}, \mathbb{R})$, the set of bounded, continuous, and measurable, real-valued functions from $\mathbf{Y}$ into $\mathbb{R}$. Let π_0 be an arbitrary initial distribution for y_0. We have

$$\begin{aligned} \lim_{t\to\infty} \int f(y_t) \pi_t\,(dy_t) &= \lim_{t\to\infty} \int f \Bigg(\Bigg(C^t \sum_{b_1} \cdots \sum_{b_t} \left(\frac{c^{b_1} \cdots c^{b_t}}{C^t} \right) y_{a+b_1+\cdots+b_t,0} \\ &\quad + \sum_{m=1}^{\infty} \sum_{b_1} \cdots \sum_{b_m} c_{b_1} \cdots c_{b_{m-1}} \Big[d_{b_s} \, \theta_{a+b_1+\cdots+b_s,1-m} \\ &\quad + \sum_{\tau=1}^{\infty} e_{b_m,\tau} \, E[\theta_{a+b_1+\ldots+b_m,1-m+\tau} \,|\, \theta_{1-m}, \ldots, \theta_{1-m-s}] \Big] \Bigg)_{a \in \mathbb{A}} \Bigg) P(d\theta) \pi_0(dy_0) \\ &= \int f \Bigg(\Bigg(\sum_{m=1}^{\infty} \sum_{b_1} \cdots \sum_{b_m} c_{b_1} \cdots c_{b_{m-1}} \Big[d_{b_s} \, \theta_{a+b_1+\cdots+b_s,1-m} \\ &\quad + \sum_{\tau=1}^{\infty} e_{b_m,\tau} \, E[\theta_{a+b_1+\ldots+b_m,1-m+\tau} \,|\, \theta_{1-m}, \ldots, \theta_{1-m-s}] \Big] \Bigg)_{a \in \mathbb{A}} \Bigg) P(d\theta) \pi_0(dy_0) \\ &= \int f\,(y)\, \pi\,(dy) \end{aligned} \tag{D.4}$$

The first equality is from (D.2); the second is due to Lebesgue Dominated Convergence theorem (see e.g. Aliprantis and Border (2006), p. 415); third is due to the continuity of f and the pointwise limit of y_t in (D.3). Thus, for any $f \in C(\mathbf{Y}, \mathbb{R})$, $\lim_{t\to\infty} \int f d\pi_t = \int f d\pi$, meaning that the sequence of equilibrium distributions π_t generated by the probability measure P and the policy g^* converges weakly to the invariant distribution π. The choice of π_0 was arbitrary. Hence, for any initial distribution, the induced equilibrium process 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 $t = 1, 2, \ldots$. From the above convergence argument $\pi_t \to \pi$ weakly. Hence $\hat{\pi} = \pi$.

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

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} f(y_t) = \int f(y_t)\pi(dy_t)$$

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

E Identification

E.1 Proof of Theorem 6: Identification

Assume first that population size N is finite. From Theorem 1, we have that if T is finite,

$$\begin{aligned} y_{a,t} &= \sum_{b\in\mathbb{A}} c_{b,T-(t-1)}\, y_{a+b,t-1} + \gamma \sum_{b\in\mathbb{A}} d_{b,T-(t-1)}\, x_{a+b,t} \\ &\quad + \gamma \sum_{\tau=t+1}^{T} \sum_{b\in\mathbb{A}} e_{b,T-(t-1),\tau-t}\, E\left[x_{a+b,\tau}|x^t\right] + \varepsilon_{a,t} \\ \varepsilon_{a,t} &= \sum_{b\in\mathbb{A}} d_{b,T-(t-1)}\, u_{a+b,t} + \sum_{\tau=t+1}^{T} \sum_{b\in\mathbb{A}} e_{b,T-(t-1),\tau-t}\, E\left[u_{a+b,\tau}|u^t\right] \end{aligned}$$

And if T is infinite,

$$\begin{aligned} y_{a,t} &= \sum_{b\in\mathbb{A}} c_b\, y_{a+b,t-1} + \gamma \sum_{b\in\mathbb{A}} d_b\, x_{a+b,t} \\ &\quad + \gamma \sum_{\tau=t+1}^{\infty} \sum_{b\in\mathbb{A}} e_{b,\tau-t} E\left[x_{a+b,\tau}|x^t\right] + \varepsilon_{a,t} \\ \varepsilon_{a,t} &= \sum_{b\in\mathbb{A}} d_b\, u_{a+b,t} + \sum_{\tau=t+1}^{\infty} \sum_{b\in\mathbb{A}} e_{b,\tau-t}\, E\left[u_{a+b,\tau}|u^t\right] \end{aligned}$$

These are econometric equations with N endogenous variables on the right hand side. Consider the N instruments $(x_{a+b,t-1})_b$. These equations can be consistently estimated through instrumental regressions if the following three sets of conditions are satisfied. (Remember that we maintain N fixed here and consider an arbitrarily large number of replications of the economy).

(1) The instruments are not perfectly correlated, which is guaranteed by Assumption 3.

(2) The appropriate exclusion restrictions are satisfied, i.e., $E\left[\varepsilon_t|x_{t-1}\right]=0$. Note that ε_t is a linear function of two sets of variables: u_t and $E\left[u_\tau|u^t\right]$, with $\tau \geq t+1$. By Assumption 2, we have $E\left[u_t|x_{t-1}\right]=0$ and, since $E\left[u_\tau|u^t\right]$ is a function of u^t, $E\left[E\left[u_\tau|u^t\right]|x_{t-1}\right]=0$.

(3) The instruments have an impact on the endogenous variables, which is the case when $\gamma \neq 0$. If $\alpha_1 \neq 0$, then $y_{a+b,t}$ is affected by $y_{a+b,s}$. If $T \geq 2$, we recover at least $(c_{b,1})_b$ and $(c_{b,2})_b$. If T is infinite, we recover $(c_b)_b$.

Suppose next that T is infinite. We know from Theorem 3-(i) that $(c_b)_{b\in\mathbb{A}}$ is injective in α_1, α_3 and β. This shows identification when N is finite. To conclude, suppose that N is infinite and denote by $\mathbb{A}_k$ the set which includes agent 0 and his k closest neighbors on the left and on the right. If T is finite, write the equilibrium characterization as follows

$$y_{a,t}=\sum_{b\in\mathbb{A}_k} c_{b,T-(t-1)}\, y_{a+b,t-1}+u_{a,t}+\varepsilon_{a,t}$$

where $\varepsilon_{a,t}$ is defined as above and $u_{a,t}=\sum_{b\in\mathbb{A}\setminus\mathbb{A}_k} c_{b,T-(t-1)}\, y_{a+b,t-1}$ and similarly for T infinite. Consider estimating these equations through instrumental regressions as above. Here, $E(u_t|y_s)\neq 0$. However, $c_{b,T-(t-1)}$ is positive for any $b\in\mathbb{A}$ and converges to zero monotonically as $|b|\to\infty$, and their sum is less than 1. This implies that $\left(\sum_{b\in\mathbb{A}\setminus\mathbb{A}_k} c_{b,T-(t-1)}\right)\to 0$ as $k\to\infty$. Hence, for a given $\epsilon>0$, there exists a k large enough s.t. $\sum_{b\in\mathbb{A}\setminus\mathbb{A}_k} c_{b,T-(t-1)} < \frac{\epsilon}{\max\{|\underline{y}|,|\bar{y}|\}}$, where $y_{a+b,t-1}\in Y=[\underline{y},\bar{y}]$ for each $b\in\mathbb{A}$. Consequently,

$$\begin{aligned}
u_{a,t} &= \sum_{b\in\mathbb{A}\setminus\mathbb{A}_k} c_{b,T-(t-1)}\, y_{a+b,t-1}\\
&\leq \left(\sum_{b\in\mathbb{A}\setminus\mathbb{A}_k} c_{b,T-(t-1)}\right)\max\{|\underline{y}|,|\bar{y}|\}\\
&< \epsilon
\end{aligned}$$

Since the choice of ϵ is arbitrary, $u_{a,t}$ becomes arbitrarily small as k tends to infinity. This implies that the difference between the estimated coefficients and their true values become arbitrarily small as k tends to infinity. Thus, if $(\alpha_1',\alpha_3',\beta')\neq(\alpha_1,\alpha_3,\beta)$, there exists k_0 such that for any $k\geq k_0$ the previous procedure is able to differentiate between outcomes generated by one set of structural coefficients vs the other. ■

E.2 Proof of the First-order Conditions

The utility of agent a at $T-1$ is equal to $v_{a,T-1}=u_{a,T-1}+\beta E u_{a,T}$ where the expectation is taken conditional on θ^{T-1} . The first-order condition is:

$$\frac{\partial u_{a,T-1}}{\partial y_{a,T-1}}+\beta\frac{\partial E u_{a,T}}{\partial y_{a,T-1}}=0$$

Compute the derivative of $v_{a,T-1}$ with respect to $y_{a,T-1}$. We get:

$$\frac{1}{2}\frac{\partial u_{a,T-1}}{\partial y_{a,T-1}}=-\alpha_1(y_{a,T-1}-y_{a,T-2})-\alpha_2(y_{a,T-1}-\theta_{a,T-1})-\alpha_3(y_{a,T-1}-y_{a-1,T-1})-\alpha_3(y_{a,T-1}-y_{a+1,T-1})$$

And

$$\begin{aligned}
\frac{1}{2}\frac{\partial Eu_{a,T}}{\partial y_{a,T-1}} &= \frac{1}{2}E\frac{\partial u_{a,T}}{\partial y_{a,T-1}} \\
&= E[-\alpha_1(y_{a,T}-y_{a,T-1})(\frac{\partial y_{a,T}}{\partial y_{a,T-1}}-1)-\alpha_2(y_{a,T}-\theta_{a,T})\frac{\partial y_{a,T}}{\partial y_{a,T-1}} \\
&- \alpha_3(y_{a,T}-y_{a-1,T})(\frac{\partial y_{a,T}}{\partial y_{a,T-1}}-\frac{\partial y_{a-1,T}}{\partial y_{a,T-1}})-\alpha_3(y_{a,T}-y_{a+1,T})(\frac{\partial y_{a,T}}{\partial y_{a,T-1}}-\frac{\partial y_{a+1,T}}{\partial y_{a,T-1}})]
\end{aligned}$$

From Theorem 1, we know that $\frac{\partial y_{a,T}}{\partial y_{a,T-1}} = c_{0,1}$ and $\frac{\partial y_{a+1,T}}{\partial y_{a,T-1}} = \frac{\partial y_{a-1,T}}{\partial y_{a,T-1}} = c_{1,1}$. We also know from the first-order conditions of the last period that

$$-\alpha_1(y_{a,T}-y_{a,T-1})-\alpha_2(y_{a,T}-\theta_{a,T})-\alpha_3(y_{a,T}-y_{a-1,T})-\alpha_3(y_{a,T}-y_{a+1,T}) = 0$$

This implies that

$$\frac{1}{2}\frac{\partial Eu_{a,T}}{\partial y_{a,T-1}} = \alpha_1(Ey_{a,T}-y_{a,T-1})+\alpha_3 c_{1,1}(2Ey_{a,T}-Ey_{a-1,T}-Ey_{a+1,T})$$

Regrouping terms, we get:

$$\begin{aligned}
(\alpha_1+\alpha_2+2\alpha_3+\beta\alpha_1)\,y_{a,T-1} &= \alpha_1\,y_{a,T-2}+\alpha_2\,\theta_{a,T-1}+\alpha_3\,(y_{a-1,T-1}+y_{a+1,T-1}) \\
&+ \beta[(\alpha_1+2\alpha_3 c_{1,1})Ey_{a,T}-\alpha_3 c_{1,1}(Ey_{a-1,T}+Ey_{a+1,T})]
\end{aligned}$$

Next, replace the expected values by their realized counterparts. This means adding an error term equal to the difference between the two. More precisely, introduce

$$\nu_{a,T-1} = \beta[(\alpha_1+2\alpha_3 c_{1,1})(Ey_{a,T}-y_{a,T})-\alpha_3 c_{1,1}(Ey_{a-1,T}-y_{a-1,T}+Ey_{a+1,T}-y_{a+1,T})]$$

Note that we have:

$$E(\nu_{a,T-1}\,|\,\theta^{T-1})=0$$

and, through the law of iterated expectations,

$$E(\nu_{a,T-1}\,|\,\theta^{s})=0$$

if $s<T-1$. Then:

$$\begin{aligned}
y_{a,T-1} &= \frac{\alpha_1}{1+\beta\alpha_1}y_{a,T-2}+\frac{\alpha_3}{1+\beta\alpha_1}(y_{a-1,T-1}+y_{a+1,T-1}) \\
&+ \frac{\beta}{1+\beta\alpha_1}\,[(\alpha_1+2\alpha_3 c_{1,1})y_{a,T}-\alpha_3 c_{1,1}(y_{a-1,T}+y_{a+1,T})] \\
&+ \frac{\alpha_2}{1+\beta\alpha_1}\theta_{a,T-1}+\frac{1}{1+\beta\alpha_1}\nu_{a,T-1}
\end{aligned}$$

This is an econometric equation expressing $y_{a,T-1}$ as a function of four endogenous variables: $y_{a,T-2}$, $y_{a-1,T-1}+y_{a+1,T-1}$, $y_{a,T}$ and $y_{a-1,T}+y_{a+1,T}$. We can simplify this further. Recall that at T, we have:

$$(\alpha_1+\alpha_2+2\alpha_3)y_{a,T} = \alpha_1\,y_{a,T-1}+\alpha_2\theta_{a,T}+\alpha_3(y_{a-1,T}+y_{a+1,T})$$

and hence

$$2\alpha_3 y_{a,T}-\alpha_3(y_{a-1,T}+y_{a+1,T}) = \alpha_1 y_{a,T-1}+\alpha_2\theta_{a,T}-(\alpha_1+\alpha_2)y_{a,T}$$

Substituting yields

$$\begin{aligned}(1+\beta\alpha_1)y_{a,T-1} &= \alpha_1 y_{a,T-2} + \alpha_2\theta_{a,T-1} + \alpha_3(y_{a-1,T-1} + y_{a+1,T-1}) \\ &+ \beta[(\alpha_1 - c_{1,1}(\alpha_1+\alpha_2))y_{a,T} + \alpha_1 c_{1,1} y_{a,T-1} + \alpha_2 c_{1,1}\theta_{a,T}] + \nu_{a,T-1}\end{aligned}$$

This yields:

$$\begin{aligned}(1+\beta\alpha_1(1-c_{1,1}))y_{a,T-1} &= \alpha_1 y_{a,T-2} + \alpha_2\theta_{a,T-1} + \alpha_3(y_{a-1,T-1} + y_{a+1,T-1}) \\ &+ \beta[\alpha_1 - c_{1,1}(\alpha_1+\alpha_2)]y_{a,T} + \beta\alpha_2 c_{1,1}\theta_{a,T} + \nu_{a,T-1}\end{aligned}$$

and

$$\begin{aligned}(1+\beta\alpha_1(1-c_{1,1}))y_{a,T-1} &= \alpha_1 y_{a,T-2} + \alpha_3(y_{a-1,T-1} + y_{a+1,T-1}) + \beta[\alpha_1 - c_{1,1}(1-2\alpha_3)]y_{a,T} \\ &+ \gamma\alpha_2 x_{a,T-1} + \gamma\beta\alpha_2 c_{1,1} x_{a,T} + \varepsilon_{a,T-1}\end{aligned}$$

where

$$\varepsilon_{a,T-1} = \alpha_2 u_{a,T-1} + \beta\alpha_2 c_{1,1} u_{a,T} + \nu_{a,T-1}$$

■

Technical Appendix to: "Dynamic Linear Economies with Social Interactions." (Not for publication!)

Onur Özgür[49]

Melbourne Business School

onur.ozgur@mbs.edu

Alberto Bisin[50]

New York University, NBER

alberto.bisin@nyu.edu

Yann Bramoullé[51]

Aix-Marseille University, CNRS

yann.bramoulle@univ-amu.fr

[49]Melbourne Business School, 200 Leicester Street, Carlton, VIC 3053, Australia; CIREQ, CIRANO

[50]Department of Economics, New York University, 19 West Fourth Street, New York, NY, 10012, USA; CIREQ, IZA, NBER.

[51]Aix-Marseille University, Centre de la Vieille Charité, 2 rue de la Charité, 13002 Marseille, France; CNRS, QREQAM.

F Lemmata used in the Existence and Uniqueness Proof

Lemma 1 (Convex Combination Form) *For any history (y^{t-1}, θ^t), the unique symmetric solution depends solely on last period equilibrium choices and current preference shock realizations, i.e. $y_1^*(y^0, \theta^1) = g_1(y_0, \theta_1)$, for some $g_1 : Y \times \Theta \to Y$. Moreover, the policy function g_1 has the convex combination form as in the statement of the theorem.*

Proof: Let

$$G := \left\{ \begin{array}{l} g : Y \times \Theta^s \to Y \text{ s.t.} \\ g(y_0,\ \theta^1) = \sum_{a\in\mathbb{A}} c_a\, y_{a,0} + \sum_{a\in\mathbb{A}} d_a\, \theta_{a,1} + \sum_{\tau=t+1}^{T} \sum_{a\in\mathbb{A}} e_{a,\tau-t}\, E\left[\theta_{a,\tau}|\theta^1\right] \\ \text{with} \\ \ \text{(i) } c_a, d_a, e_a \geq 0 \text{ and } \sum_{a\in\mathbb{A}} \left(c_a + d_a + \sum_{\tau=t+1}^{T} e_{a,\tau-t}\right) = 1 \\ \ \text{(ii) } (\tfrac{1}{2})c_{a+1} + (\tfrac{1}{2})c_{a-1} \geq c_a, \forall a \neq 0 \\ \ \text{(iii) } c_b \leq c_a, \forall a, b \in \mathbb{A} \text{ with } |b| > |a|. \\ \ \text{(iv) } c_a = c_{-a},\ \forall a \in \mathbb{A} \\ \text{and properties (ii), (iii), and (iv) also holding for the } d \text{ and } e \text{ sequences.} \end{array} \right\} \tag{F.1}$$

be the class of functions that are convex combinations (i) of one-period before history, current and expected future preference shocks, having the (ii) 'convexity', (iii) 'monotonicity', and (iv) 'symmetry' properties. Let $g \in G$ be such that after any history $(y^0, \theta^1) = (y_{-(s-1)}, \theta_{-(s-2)}, \ldots, Y_{-1}, \theta_0, y_0, \theta_1)$

$$y_1(y^0, \theta^1) = g(y_0, \theta_1)$$

and let (c, d, e) be the coefficient sequence associated with g. Applying L_1 to y_1 (hence to g), we get

$$\begin{aligned} (L_1 y_1)\left(y^0, \theta^1\right) &= \Delta_1^{-1}\left(\alpha_1\, y_{0,0} + \alpha_2\, \theta_{0,1} + \alpha_3\, g\left(R^{-1} y_0, R^{-1}\theta_1\right) + \alpha_3\, g\left(R\, y_0, R\, \theta_1\right)\right) \\ &= \Delta_1^{-1}\Bigg[\alpha_1\, y_{0,0} + \alpha_2\, \theta_{0,1} + \alpha_3\left(\sum_{a\in\mathbb{A}} c_a\, y_{a-1,0} + \sum_{a\in\mathbb{A}} d_a\, \theta_{a-1,1}\right) \\ &\qquad +\alpha_3\left(\sum_{a\in\mathbb{A}} c_a\, y_{a+1,0} + \sum_{a\in\mathbb{A}} d_a\, \theta_{a+1,1}\right)\Bigg] \end{aligned} \tag{F.2}$$

By the definition of G in (F.1), the coefficient sequences are positive and absolutely summable; the choices and shocks are elements of a compact set. Hence, we can rearrange the series to obtain

$$\begin{aligned} = \ \Delta_1^{-1}\Bigg(&\underbrace{(\alpha_1 + \alpha_3\, c_{-1} + \alpha_3\, c_1)}_{\Delta_1 c_0'}\, y_{0,0} + \underbrace{(\alpha_2 + \alpha_3\, d_{-1} + \alpha_3\, d_1)}_{\Delta_1 d_0'}\, \theta_{0,1} \\ &+ \sum_{a\neq 0} \underbrace{(\alpha_3\, c_{a-1} + \alpha_3\, c_{a+1})}_{\Delta_1 c_a'}\, y_{a,0} + \sum_{a\neq 0} \underbrace{(\alpha_3\, d_{a-1} + \alpha_3\, d_{a+1}}_{\Delta_1 d_a'}\, \theta_{a,1}\Bigg) \end{aligned} \tag{F.3}$$

This last expression is linear in y_0, and θ_1. So, $L_1 y_1$ preserves the same linear form. By definition of the new coefficient sequence (c', d') in (F.3), 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 (F.2) equal to 1. Thus, the total sum of coefficients on the right hand side of (F.2) is $\Delta_1^{-1}(\alpha_1+\alpha_2+2\alpha_3)=1$, which proves property (i). The final form in (F.3) is just a regrouping of elements in (F.2). Let $(c'_a)_{a\in\mathbb{A}}$ be the new coefficient sequence associated with $L_1 y_T$ as defined in equation (F.3). Pick $a\neq 0$ in $\mathbb{A}$,

$$\begin{aligned}
c'_{a+1}+c'_{a-1} &\geq \left(\frac{\alpha_3}{\Delta_1}\right)(c_a+c_{a+2})+\left(\frac{\alpha_3}{\Delta_1}\right)(c_{a-2}+c_a)\\
&\geq \left(\frac{\alpha_3}{\Delta_1}\right)(2c_{a+1}+2c_{a-1})\\
&= 2\left(\frac{\alpha_3}{\Delta_1}\right)(c_{a+1}+c_{a-1})\\
&= 2c'_a
\end{aligned}$$

By definition of c' in (F.3), 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 c' in (F.3). Therefore, for any $a\neq 0$, $c_{a+1}+c'_{a-1}\geq 2c'_a$, which is property (ii). Now, pick any $a,b\in\mathbb{A}$ with $|a|<|b|$.

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

First equality is from (F.3); second by property (iv) of G in (F.1); the inequality is property (iii) of G in (F.1); next equality is due to property (iv) of G again; and finally the last equality is by (F.3). Hence, property (iii) in (F.1) holds for the new sequence. We next show that c' satisfies (iv) in (F.1).

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

where first equality is by (F.3); the second is due to (iv) of G in (F.1); finally the last is again by (F.3).

Thus, the restriction of L_1 to the subspace (call it B_G) of bounded measurable functions that agree with an element of G after any history, maps elements of B_G into itself. Moreover, endowed with the sup norm, B_G is a closed subset of $B\left((Y\times\Theta)^s,Y\right)$ 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 as well and the unique fixed point y_1^* in $B\left((Y\times\Theta)^s,Y\right)$ 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 assume the convex combination form stated in the theorem. This concludes the proof of Lemma 1. ■

Lemma 2 (Interiority) *Let $T\geq 2$. The unique optimizer $y_{a,1}$ is almost surely in the interior of $Y=[\underline{y},\,\bar{y}]$, and equation (A.5) can be written as*

$$0 \;=\; -y_{a,1}\,\Delta_T+\alpha_1\,y_{a,0}+\alpha_2\,\theta_{a,1}+\sum_{b\neq 0}\gamma_{b,T}\,y_{a+b,1}+\sum_{b\in\mathbb{A}}\sum_{\tau=2}^{T}\mu_{b,\tau,T}\,E\left[\theta_{a+b,\tau}|\theta^t\right] \qquad \text{(F.4)}$$

where $\Delta_T := \alpha_1+\alpha_2+\sum_{b\neq 0}\gamma_{b,T}+\sum_{b\in\mathbb{A}}\sum_{\tau=2}^{T}\mu_{b,\tau,T}$, *and the coefficients* α_1, α_2, $(\gamma_{b,T})_{b\neq 0}$, *and* $(\mu_{b,\tau,T})_{b\in\mathbb{A}}^{\tau\geq 2}$ *are non-negative.*

Proof: In order to prove interiority, we let

$$\tau := \inf\left\{t>0:\ g(y_{t-1},\theta^t)=\underline{y}\right\} \quad \text{and} \quad y_t := g(y_{t-1},\theta^t).$$

It suffices to show that $Prob[\tau<T]=0$. Let us assume to the contrary that $Prob[\tau<T]>0$. In such a situation, $y_\tau=\underline{y}$ is optimal and a necessary condition for optimality is

$$\begin{aligned} & -\alpha_1(\underline{y}-y_{a,\tau-1})^2-\alpha_2(\underline{y}-\theta_{a,\tau})^2-\alpha_3(\underline{y}-y_{a-1,\tau})^2-\alpha_3(\underline{y}-y_{a+1,\tau})^2-\beta\alpha_1 E[(\underline{y}-y_{a,\tau+1})^2\,|\,\theta^\tau] \\ \geq\ & -\alpha_1(y-y_{a,\tau-1})^2-\alpha_2(y-\theta_{a,\tau})^2-\alpha_3(y-y_{a-1,\tau})^2-\alpha_3(y-y_{a+1,\tau})^2-\beta\alpha_1 E[(y-y_{a,\tau+1})^2\,|\,\theta^\tau] \end{aligned}$$

for all $y\in Y$, since otherwise $y_\tau<\underline{y}$ would lead to a higher payoff. This, however, requires $\theta_{a,\tau}=y_{a,\tau-1}=y_{a,\tau+1}=\underline{y}$. This shows that $y_{a,t}=\underline{y}=\theta_{a,t}$ for all $t\leq T$. This, of course, contradicts $E[\theta_{a,t}\,|\,\theta^{t-1}]\in(\underline{y},\bar{y})$. Thus, $Prob[\tau<T]=0$.

Moreover, thanks to the linearity of the first order condition in the choice variables, the preference shocks, and the expected future preference shocks, one can write the first order condition in equation (A.5) as a function only of contemporaneous choices, and expected future shocks, through iterative application of the policy functions for future period equilibrium choices, as we demonstrated in (A.7). Finally, since the unique optimizer $y_{a,1}$ is almost surely interior, the coefficients multiplying these are necessarily non-negative. Finally, since all values inside the brackets are uniformly bounded and the finite horizon equilibria converge to the infinite horizon equilibrium uniformly, all statements hold for $T=\infty$ as well. This concludes the proof of the Lemma. ■

Lemma 3 (Compactness) *L_β and G endowed with the supnorm are compact metric spaces.*

Proof: Let $(\beta_{T_n})_n$ be a sequence lying in L_β that converges to $y=(y_t)\in[0,1]^\infty$. This means that $\beta_{T_n,t}\to y_t$, for all $t\geq 1$, which in turn means that $y_t\in\{0,\beta^t\}$ by the construction of L_β. Moreover, if $y_t=0$ for some t, $y_{t+\tau}=0$ for all $\tau\geq 1$ since the terms β_{T_n} are geometric (finite or infinite) sequences. There are two possibilities: either $y=(1,\beta,\ldots,\beta^T,0,0,\ldots)$ or $y=\beta^t$ 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 a natural number $N\geq 1$ 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 $\{\beta_1,\beta_2,\ldots,\beta_N\}\subset L_\beta$. 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 $H := \left\{y=(y_a)_{a\in\mathbb{A}}\mid y_a\leq\left(\frac{1}{2a}\right),\ \text{for all } a\in\mathbb{A}\right\}$. 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 $\bar{y}\in H$. For any $a\in\mathbb{A}$ s.t. $|a|\geq N$, $[0,(2N)^{-1}]\subset B_\infty(y_a,\epsilon)$, the ϵ-ball around y_a with respect to the sup norm. For $|a|\leq N$, let $Y(a):=\left\{0,\epsilon,2\epsilon,\ldots,k_a\epsilon,(2a)^{-1}\right\}$, where k_a is the greatest integer s.t. $k_a\epsilon\leq(2a)^{-1}$. The set

$$\left\{y\in H\ \middle|\ y_a=\bar{y}_a,\ \text{for } |a|\geq N,\ \text{and } (y_{-(N-1)},\ldots,y_0,\ldots,y_{N-1})\in\prod_{|a|\leq N}Y(a),\ \text{for } |a|\leq N\right\}$$

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 $((c_a, d_a, e_a)_a)$. Clearly, for any sequence of policies in G, $g_n \to g$ in sup norm if and only if the associated coefficients $((c_{a,n}, d_{a,n}, e_{a,n})_a) \to ((c_a, d_a, e_a)_a)$ in sup norm. We know from (A.3) that c satisfies properties $(i), (ii)$ and (iii). Thus, for any $a \in \mathbb{A}$, $c_0 > c_1 > \ldots > c_{|a|}$, $c_a = c_{-a}$ 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 and $e = (e_{b,\tau})_{b \in \mathbb{A}}^{\tau \geq 1}$ 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 $(c_{m_n}, d_{m_n}, e_{m_n}) \to (c, d, e) \in L_G$. The latter, being an admissible coefficient sequence, is associated with the policy $g(y, \theta^s) := \sum_a c_a y_a + \sum_a d_a \theta_{a,s} + \sum_a \sum_{\tau \geq 1} e_{a,\tau} \, E\left[\theta_{a,s+\tau} | \theta^s\right]$. Thus, the respective policy subsequence $g_{m_n} \to g \in G$. This establishes that G is sequentially compact hence compact. This concludes the proof of Lemma 3. ■

Lemma 4 (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.*

Proof: Since G endowed with the sup norm is a compact metric space due to Lemma 3, 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 Y is compact. Thus, one can choose a $\delta' > 0$ such that for any t, $|x_{0,t} - y_{0,t}| < \delta'$ implies

$$\left| u\left(x_{0,t-1}, x_{0,t}, \{x_{b,t}(g)\}_{b \in \{-1,1\}}, \theta_{0,t}\right) - u\left(y_{0,t-1}, y_{0,t}, \{x_{b,t}(g)\}_{b \in \{-1,1\}}, \theta_{0,t}\right)\right| \; < \; \epsilon'.$$

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 $t \leq T$, $||g^0_t - g'^0_t||_\infty < 2^T \delta = \delta'$ hence $|y_{0,t}(g^0) - y_{0,t}(g'^0)| < \delta$. 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

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

■

G Convergence Properties of Policy Coefficient Sequences

In this section, we study the dynamic and cross-sectional properties of the equilibrium coefficient sequences in some detail. Throughout the section, we present equilibrium arguments using c_T. Arguments for d_T and e_T are identical. To do that, we index coefficients by the length T of the economy, i.e., $c_T = (c_{a,T})_{a\in\mathbb{A}}$ is the vector of policy coefficients on "history" for the optimal choice in the first period of an economy with T periods. For $T = \infty$, we drop the time index and simply write $c = (c_a)_{a\in\mathbb{A}}$.

We first present the statements of the three Lemmas we will prove in this section. Lemma 5 is used in the proof of Theorem 2. Lemma 6 is used in the proofs of Theorem 2 and Theorem 3. Finally, Lemma 7 is used in the proof of Theorem 3. We then provide the proofs of these results in the subsection G.1 that follows.

We know from Theorem 1 that the policy coefficients form non-negative, bounded, absolutely summable sequences. More specifically, for any $T \geq 1$, any $a \in \mathbb{A}$, the coefficients $c_{b,T}$ satisfy

$$\lim_{|b|\to\infty} c_{a+b,T} = 0$$

The impact of an agent $a+b$ on agent a tends to zero as $|b| \to \infty$. Furthermore, equilibrium policy functions are non-stationary in the finite economy, as rational forward-looking agents change their behaviour optimally through time. Finally, the finite-horizon parameters converge (uniformly) to the infinite-horizon stationary policy parameters,

$$\lim_{T\to\infty} c_T = c$$

To study these convergence properties precisely, we define the **cross-sectional rates of convergence** as

$$r_{|a|+1,T} \quad := \quad \frac{c_{|a|+1}}{c_{|a|,T}}, \qquad \text{for any } a \in \mathbb{A}$$

From Lemma 2, the first order condition (A.5) characterizing any agent a's optimal T-period choice can be written in a more concise way as in (A.10). So, by matching equilibrium coefficients c_T on both sides of (A.10), one obtains for any agent $a \in \mathbb{A}$

$$c_{a,T} \quad = \quad \Delta_T^{-1}\left[\alpha_1\, I_{\{a=0\}} + \sum_{b\neq 0}\gamma_{b,T}\, c_{a-b,T}\right] \tag{G.1}$$

where $\gamma_{b,T}$ is the quantified impact on expected discounted marginal utility of agent a, of a change in individual $a+b$'s first period choice, $y_{a+b,1}$, as defined in (A.9) in the proof of Lemma 2 Formally, the expression $\sum_{b\neq 0}\gamma_{b,T}\, c_{a-b,T}$ inside the brackets in equation (G.1) is the discrete convolution of the policy coefficient sequence $c_T = (c_{a,T})_{a\in\mathbb{A}}$ and the coefficient sequence $\gamma_{b,T} = (\gamma_{b,T})_{b\neq 0}$, where a acts as the shift parameter.

Our first result characterizes the monotonicity of the cross-sectional convergence rates in the parameter α_3 and in social distance.

Lemma 5 (Monotone Increasing Cross-Sectional Rates) *For any $T \geq 2$, the following hold:*

(i) The rates at which the policy coefficients converge to zero at the cross-section are strictly monotonically increasing in $|a|$, i.e., for any $a \in \mathbb{A}$

$$r_{|a|+1,T} = \frac{c_{|a|+1,T}}{c_{|a|,T}} \quad < \quad \frac{c_{|a|+2,T}}{c_{|a|+1,T}} = r_{|a|+2,T} \tag{G.2}$$

(ii) Given β and $\frac{\alpha_1}{\alpha_1+\alpha_2}$, the cross-sectional rates are strictly increasing in α_3, i.e.,

$$r_{a,T}(\alpha_3') \quad > \quad r_{a,T}(\alpha_3), \qquad \text{for any } a \neq 0. \tag{G.3}$$

The analogous results hold for for d_T and e_T.

Our next result characterizes the behaviour of the **sums of policy coefficients across periods**. We rely on these results in the identification proofs. In order to demonstrate it formally, we first define the sum of policy coefficients on "history", "current shocks" (own effect), and on "expectations" respectively as

$$\begin{aligned} C_T &:= \sum_b c_{b,T} \\ D_T &:= \sum_b d_{b,T} \\ E_T &:= \sum_{\tau=t+1}^{T} \sum_b e_{b,T,\tau-t} \end{aligned}$$

where $C_T + D_T + E_T = 1$. The result we will present is interesting also in the sense that the sums follow a particular recursive structure well-known and extremely useful in mathematics and dynamic systems: They behave as **continued fractions**. A continued fraction is an expression obtained through an iterative process of representing a number as the sum of its integer part and the reciprocal of another number, then writing this other number as the sum of its integer part and another reciprocal, and so on. In a finite continued fraction, the iteration/recursion is terminated after finitely many steps. In contrast, an infinite continued fraction is an infinite expression. In either case, all integers in the sequence, other than the first, must be positive. The integers are called the coefficients or terms of the continued fraction (see e.g. Pettofrezzo and Byrkit (1970)). In our environment, they take the following form

$$C_T = \cfrac{\alpha_1}{(\alpha_1+\alpha_2+\alpha_1\beta) - \beta\cfrac{\alpha_1}{(\alpha_1+\alpha_2+\alpha_1\beta) - \beta\cfrac{\alpha_1}{(\alpha_1+\alpha_2+\alpha_1\beta) - \beta\cfrac{\alpha_1}{\cdots}}}}$$

and the next Lemma characterizes their recursive properties and their limit behaviour as the number of periods, T, increases arbitrarily.

Lemma 6 (Policy Coefficient Sums) *For a T-period dynamic conformity economy with $T > 1$, the policy coefficient sums for $l = 2, \ldots, T$ are given by the following recursive system of continued fractions*

$$\begin{aligned} C_l &= \frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_1\beta\,(1-C_{l-1})} \\ D_l &= \frac{\alpha_2}{\alpha_1+\alpha_2+\alpha_1\beta\,(1-C_{l-1})} \\ E_l &= \frac{\alpha_1\beta\,(1-C_{l-1})}{\alpha_1+\alpha_2+\alpha_1\beta\,(1-C_{l-1})} \end{aligned} \tag{G.4}$$

where $C_1 = \frac{\alpha_1}{\alpha_1+\alpha_2}$*,* $D_1 = \frac{\alpha_2}{\alpha_1+\alpha_2}$*, and* $E_1 = 0$*. Moreover,* $C_l \downarrow C_\infty$ *and* $D_l \downarrow D_\infty$ *are monotonically decreasing (hence* $E_l \uparrow E_\infty$*) sequences where* C_∞*,* D_∞*, and* E_∞ *are the fixed points of the respective equations in the recursive system (G.4).*

Finally, our final lemma, Lemma 7, shows that there are at least two agents, b and b', for whom the marginal rates of substitution between α_1 and p, $MRS^b_{\alpha_1,p}$ and $MRS^{b'}_{\alpha_1,p}$, that maintain the same level for the maps $\frac{\gamma_b(\alpha_1,p,\beta(p))}{\alpha_1}$ and $\frac{\gamma_{b'}(\alpha_1,p,\beta(p))}{\alpha_1}$, are different at the true parameter pair (where the level curves intersect). This variation in the responsiveness of the level sets of $\frac{\gamma_b(\alpha_1,p,\beta(p))}{\alpha_1}$ to changes in α_1 and p, as b spans the cross-section, helps us identify the true parameter pair (α_1, p) consistent with observed levels for the policy coefficients $(c_a)_{a\in\mathbb{A}}$.

Lemma 7 (Single Crossing - Cross-section) *There exists a unique sequence* $(\bar{\gamma}_b) := (\alpha_1^{-1}\gamma_{b_1})$ *that satisfies the system in (C.3). Moreover,* (i) *the map* $\frac{\gamma_b(\alpha_1,p,\beta(p))}{\alpha_1}$ *is continuously differentiable and the partial derivatives satisfy* $\frac{\partial}{\partial p}\frac{\gamma_b(\alpha_1,p,\beta(p))}{\alpha_1} > 0$ *and* $\frac{\partial}{\partial \alpha_1}\frac{\gamma_b(\alpha_1,p,\beta(p))}{\alpha_1} < 0$*, for any* $(\alpha_1, p, \beta(p))$*; and* (ii) *there exist agents* $b \neq b'$ *such that for any* $(\alpha_1, p, \beta(p))$*,*

$$-\frac{\frac{\partial \alpha_1^{-1}\gamma_{b'}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b'}}{\partial \alpha_1}}\Bigg|_{\alpha_1^{-1}\gamma_{b'}(\alpha_1,p,\beta(p))=\bar{\gamma}'_b} > -\frac{\frac{\partial \alpha_1^{-1}\gamma_{b}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b}}{\partial \alpha_1}}\Bigg|_{\alpha_1^{-1}\gamma_{b}(\alpha_1,p,\beta(p))=\bar{\gamma}_b} > 0. \tag{G.5}$$

G.1 Proofs of Lemmata

Proof of Lemma 5: We know from Theorem 1 that there exists a unique coefficient sequence satisfying equation (G.1) and which lies in the space G of policy coefficient sequences having the desired equilibrium properties of convexity, symmetry, and monotonicity, as defined in (A.3) in the proof of Theorem 1. The right hand side of equation (G.1) maps sequences (c_b) into sequences (c'_b). We would like to show that it maps the closed subset $G_m \subset G$ into itself, where elements of G_m possess the nice properties stated in the Lemma. This would in turn imply that the policy coefficients of the unique equilibrium should lie in G_m, and should possess the properties associated with G_m.

Proof of the first part. Let $G_m \subset G$ be defined as $G_m := \{(c_a) \in G \mid \frac{c_{|a|+1}}{c_{|a|}}$ is weakly increasing in $|a|.\}$. We want to show that for any $(c_{a,T}) \in G_m$, $(c'_{a,T}) \in G_m$ as well, where $(c'_{a,T})$ is defined, using the right hand side of (G.1), as

$$c'_{a,T} \quad := \quad \Delta_T^{-1}\left[\alpha_1\, I_{\{a\in\{-1,1\}\}} + \sum_{b\neq 0}\gamma_{b,T}\, c_{a-b,T}\right] \tag{G.6}$$

G_m being defined by inequality constraints, is a closed subset of G. The map on the right hand side of (G.6) is a contraction, as we demonstrated in the proof of Theorem 1. The rest of the proof is by induction on the number of periods T.

- *Assume that* $T = 2$. We know from (A.9) in the proof of Theorem 1 that $\gamma_{b,T}$ is defined for any $b \neq 0$ as

$$\begin{aligned}\gamma_{b,T} &= \alpha_3\, I_{\{b\in\{-1,1\}\}} \\ &-\sum_{\tau=2}^{T}\beta^{\tau-1}\Bigg(\alpha_1\,\frac{\partial}{\partial y_{a+b,1}}\,(y_{a,\tau-1}-y_{a,\tau})\,\frac{\partial}{\partial y_{a,1}}\,(y_{a,\tau-1}-y_{a,\tau})+\alpha_2\,\frac{\partial}{\partial y_{a+b,1}}\,y_{a,\tau}\,\frac{\partial}{\partial y_{a,1}}\,y_{a,\tau} \\ &+\alpha_3\,\frac{\partial}{\partial y_{a+b,1}}\,(y_{a-1,\tau}-y_{a,\tau})\,\frac{\partial}{\partial y_{a,1}}\,(y_{a-1,\tau}-y_{a,\tau})+\alpha_3\,\frac{\partial}{\partial y_{a+b,1}}\,(y_{a+1,\tau}-y_{a,\tau})\,\frac{\partial}{\partial y_{a,1}}\,(y_{a+1,\tau}-y_{a,\tau})\Bigg)\end{aligned} \tag{G.7}$$

Each $\gamma_{b,T} > 0$ and $\sum_{b\neq 0}\frac{\gamma_{b,T}}{\Delta_T} < 1$, which makes the right hand side of (G.6) a contraction mapping. Using the definition in (G.7) for $T = 2$ and substituting the policy function for future choices, one obtains

$$\begin{aligned}\gamma_{b,2} &= \alpha_3\, I_{\{b\in\{-1,1\}\}} - \beta\,(\alpha_1\,(-c_{b,1})\,(1-c_{0,1}) + \alpha_2\, c_{b,1}\, c_{0,1} \\ &\quad + \alpha_3\,\,(c_{b+1,1}-c_{b,1})\,(c_{1,1}-c_{0,1}) + \alpha_3\,\,(c_{b-1,1}-c_{b,1})\,(c_{-1,1}-c_{0,1})) \\ &= \alpha_3\, I_{\{b\in\{-1,1\}\}} + \beta\, c_{b,1}\Big(\alpha_1\,\,(1-c_{0,1}) - \alpha_2\, c_{0,1} - (1-2\alpha_3)\,(c_{1,1}-c_{0,1})\Big) \\ &= \alpha_3\, I_{\{b\in\{-1,1\}\}} + \beta\, c_{b,1}\Big(\alpha_1 - (\alpha_1+\alpha_2)\, c_{0,1}\, r_1\Big)\end{aligned} \tag{G.8}$$

where the final equality uses the structure of the coefficients for $l = 1$, which we proved in Theorem 2-(ii) (see also footnote 19), and the symmetry of c_1 around zero. Remember that r_1 is the rate of convergence of the exponentially declining policy sequence for $l = 1$. We will need the convergence rates of γ_l sequences as well in the convolution. So, we define analogously, for any $l \geq 1$,

$$r^{\gamma}_{|b|+1,l} \;:=\; \frac{\gamma_{|b|+1,l}}{\gamma_{|b|,l}}, \qquad \text{for any } b \neq 0$$

Now, for $b \neq 0$, using (G.8)

$$r^{\gamma}_{|b|+1,2} \;=\; \begin{cases} \dfrac{\beta\, c_{2,1}\Big(\alpha_1-(\alpha_1+\alpha_2)c_{0,1}\, r_1\Big)}{\alpha_3+\beta\, c_{1,1}\Big(\alpha_1-(\alpha_1+\alpha_2)c_{0,1}\, r_1\Big)} \;<\; r_1, & \text{if } b\in\{-1,1\} \\ r_1, & \text{otherwise.}\end{cases} \tag{G.9}$$

Hence, $r^{\gamma}_{|b|+1,2} \geq r^{\gamma}_{|b|,2}$, for any $b \neq 0$, meaning that $r^{\gamma}_{|b|,2}$ is weakly monotonically increasing in $|b|$. Now, pick $a > 0$ wlog (the proof for $a < 0$ is identical thanks to the symmetry of the environment). From (G.1)

$$\begin{aligned}\Delta_2\, c'_{a+1,2} &= \sum_{b\neq 0}\gamma_{b,2}\, c_{a+1-b,2} \\ &= [\gamma_{2,2}c_{a-1,2} + \gamma_{3,2}c_{a-2,2} + \ldots] + [\gamma_{1,2}c_{a,2} + \gamma_{-1,2}c_{a+2,2} + \gamma_{-2,2}c_{a+3,2} + \ldots] \\ &= \left[(\gamma_{1,2}c_{a-1,2})\left(\frac{\gamma_{2,2}}{\gamma_{1,2}}\right) + (\gamma_{2,2}c_{a-2,2})\left(\frac{\gamma_{3,2}}{\gamma_{2,2}}\right) + \ldots\right] \\ &+ \left[(\gamma_{1,2}c_{a-1,2})\left(\frac{c_{a,2}}{c_{a-1,2}}\right) + (\gamma_{-1,2}c_{a+1,2})\left(\frac{c_{a+2,2}}{c_{a+1,2}}\right) + (\gamma_{-2,2}c_{a+2,2})\left(\frac{c_{a+3,2}}{c_{a+2,2}}\right) + \ldots\right] \\ &= [(\gamma_{1,2}c_{a-1,2})\, r^{\gamma}_{2,2} + (\gamma_{2,2}c_{a-2,2})\, r^{\gamma}_{3,2} + \ldots] \\ &+ [(\gamma_{1,2}c_{a-1,2})\, r_{a,2} + (\gamma_{-1,2}c_{a+1,2})\, r_{a+2,2} + (\gamma_{-2,2}c_{a+2,2})\, r_{a+3,2} + \ldots] \\ &= r'_{a+1,2}\sum_{b\neq 0}\gamma_{b,2}\, c_{a-b,2} \\ &= \Delta_2\, r'_{a+1,2}\, c'_{a,2}\end{aligned} \tag{G.10}$$

where $r'_{a+1,2}$ is the probability-weighted average rate of decline, if one interprets $\left(\frac{\gamma_{b,2}\, c_{a-b,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a-b,2}}\right)_{b\in\mathbb{A}}$ as a probability distribution on convergence rates. Similar argument for $c'_{a+2,2}$ would yield

$$\begin{aligned}
\Delta_2\, c'_{a+2,2} &= \sum_{b\neq 0}\gamma_{b,2}\, c_{a+2-b,2} \\
&= [\gamma_{3,2}c_{a-1,2}+\gamma_{4,2}c_{a-2,2}+\ \ldots]+\gamma_{2,2}c_{a,2}+[\gamma_{1,2}c_{a+1,2}+\gamma_{-1,2}c_{a+3,2}+\gamma_{-2,2}c_{a+4,2}+\ldots] \\
&= \left[(\gamma_{2,2}c_{a-1,2})\, r^{\gamma}_{3,1}+(\gamma_{3,2}c_{a-2,2})\, r^{\gamma}_{4,2}+\ldots\right]+\gamma_{2,2}c_{a,2} \qquad \text{(G.11)}\\
&+ [(\gamma_{1,2}c_{a,2})\, r_{a+1,2}+(\gamma_{-1,2}c_{a+2,2})\, r_{a+3,2}+(\gamma_{-2,2}c_{a+3,2})\, r_{a+4,2}+\ldots] \\
&= r'_{a+2,2}\sum_{b\neq 0}\gamma_{b,2}\, c_{a+1-b,2} \\
&= \Delta_2\, r'_{a+2,2}\, c'_{a+1,2}
\end{aligned}$$

where $r'_{a+2,2}$ is the probability-weighted average rate of decline, if one interprets $\left(\frac{\gamma_{b,2}\, c_{a+1-b,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a+1-b,2}}\right)_{b\in\mathbb{A}}$ as a probability distribution on convergence rates. Now, comparing the first brackets of the expressions in (G.10) and (G.11), we see that

$$\frac{\left(\frac{\gamma_{2,2}c_{a-1,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a+1-b,2}}\right)}{\left(\frac{\gamma_{1,2}c_{a-1,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a-b,2}}\right)} \leq \frac{\left(\frac{\gamma_{3,2}c_{a-2,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a+1-b,2}}\right)}{\left(\frac{\gamma_{2,2}c_{a-2,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a-b,2}}\right)} \leq \ldots \qquad \text{(G.12)}$$

due to the monotonicity of the convergence rates for γ_2, i.e., $r^{\gamma}_{2,2}=\frac{\gamma_{2,2}}{\gamma_{1,2}}\leq\frac{\gamma_{3,2}}{\gamma_{2,2}}=r^{\gamma}_{3,2}\leq\ldots$. This means that the likelihood associated with $a+1$ in (G.11) puts relatively more weight on the left tail, towards higher convergence rates for γ_2, than does the likelihood associated with a in (G.10). Similar argument applied to the expressions in the second brackets in (G.10) and (G.11) yields

$$\frac{\left(\frac{\gamma_{1,2}c_{a,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a+1-b,2}}\right)}{\left(\frac{\gamma_{1,2}c_{a-1,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a-b,2}}\right)} \leq \frac{\left(\frac{\gamma_{-1,2}c_{a+2,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a+1-b,2}}\right)}{\left(\frac{\gamma_{-1,2}c_{a+1,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a-b,2}}\right)} \leq \frac{\left(\frac{\gamma_{-2,2}c_{a+3,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a+1-b,2}}\right)}{\left(\frac{\gamma_{-2,2}c_{a+2,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a-b,2}}\right)} \leq \ldots \qquad \text{(G.13)}$$

due to the monotonicity of the convergence rates for c_2, by hypothesis, i.e., $r_{a,2}=\frac{c_{a,2}}{c_{a-1,2}}\leq\frac{c_{a+2,2}}{c_{a+1,2}}=r_{a+2,2}\leq\frac{c_{a+3,2}}{c_{a+2,2}}=r_{a+3,2}\leq\ldots$. This means that the likelihood associated with $a+1$ puts relatively more weight on the right tail, towards higher convergence rates for c_2, than does the likelihood associated with a. Moreover, each element of the likelihood sequence inside the first brackets in (G.11) multiplies a higher convergence rate than the associated member of the likelihood sequence inside the first brackets in (G.10), since $r^{\gamma}_{a+1,2}\geq r^{\gamma}_{a,2}$ for any $a\neq 0$. Similarly, each element of the likelihood sequence inside the second brackets in (G.11) multiplies a higher convergence rate than the associated member of the likelihood sequence inside the second brackets in (G.10), since $r_{a+1,2}\geq r_{a,2}$ for any $a\neq 0$. These facts combined with the monotone likelihood ratio property on each tail for the two distributions that we demonstrated above in (G.12) and (G.13) imply that the average convergence rate, $r'_{a+2,2}$, computed under the distribution $\left(\frac{\gamma_{b,2}\, c_{a+1-b,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a+1-b,2}}\right)_{b\in\mathbb{A}}$ is higher than the average convergence rate, $r'_{a+1,2}$, computed under the distribution $\left(\frac{\gamma_{b,2}\, c_{a-b,2}}{\sum_{b\neq 0}\gamma_{b,2}\, c_{a-b,2}}\right)_{b\in\mathbb{A}}$. Furthermore, the expression in (G.11) contains an extra non-zero term, $\gamma_{2,2}c_{a,2}$, which means that the ordering is strict, i.e., $r'_{a+2,2}>r'_{a+1,2}$.

So far, we established that for $T=2$, c_2 sequence satisfies the monotone increasing cross-sectional rates property; and that γ_2 satisfies this property weakly.

- *Induction Step.* Now, assume for induction that $(\gamma_l)_{l=2}^{T-1}$ satisfies the weak version and $(c_l)_{l=2}^{T-1}$ satisfies the strict version up to a period $T > 2$.

Let $u(t) := u\left(y_{0,t-1}, y_{0,t}, \{y_{b,t}\}_{b\in\{-1,1\}}, \theta_{0,t}\right)$ where u represents the conformity preferences in Assumption 1, with . Let $u_0(t) := \frac{\partial}{\partial y_{0,1}} u(t)$. From its definition in (G.7), $\gamma_{b,T}$ can be written parsimoniously as

$$\begin{aligned}\gamma_{b,T} \quad := \quad & \alpha_3\, I_{\{b\in\{-1,1\}\}} \\ & + \sum_{\tau=2}^{T} \beta^{\tau-1} \left[\left(\frac{\partial y_{0,\tau-1}}{\partial y_{b,1}}\right) \frac{\partial}{\partial y_{0,\tau-1}} u_0(\tau) + \left(\frac{\partial y_{0,\tau}}{\partial y_{b,1}}\right) \frac{\partial}{\partial y_{0,\tau}} u_0(\tau) \right. \\ & + \left. \left(\frac{\partial y_{-1,\tau}}{\partial y_{b,1}}\right) \frac{\partial}{\partial y_{-1,\tau}} u_0(\tau) + \left(\frac{\partial y_{1,\tau}}{\partial y_{b,1}}\right) \frac{\partial}{\partial y_{1,\tau}} u_0(\tau) \right]\end{aligned} \tag{G.14}$$

We will present the argument for one of the terms inside the summand. As it will be apparent, the method of proof applies to the remaining terms straightforwardly. Assume w.l.o.g. that $a \geq 0$.

$$\begin{aligned}\left(\frac{\partial y_{0,\tau}}{\partial y_{b,1}}\right) \frac{\partial}{\partial y_{0,\tau}} u_0(\tau) \quad &= \quad \sum_{s\in\mathbb{A}} \left(\frac{\partial y_{s,2}}{\partial y_{b,1}}\right) \left(\frac{\partial y_{0,\tau}}{\partial y_{s,2}}\right) \frac{\partial}{\partial y_{0,\tau}} u_0(\tau) \\ &= \quad \frac{\partial}{\partial y_{0,\tau}} u_0(\tau) \sum_{s\in\mathbb{A}} c_{b-s,T-1} \left(\frac{\partial y_{0,\tau}}{\partial y_{s,2}}\right)\end{aligned} \tag{G.15}$$

and the corresponding term for $\gamma_{b+1,T}$ is

$$\frac{\partial}{\partial y_{0,\tau}} u_0(\tau) \sum_{s\in\mathbb{A}} c_{b+1-s,T-1} \left(\frac{\partial y_{0,\tau}}{\partial y_{s,2}}\right) \tag{G.16}$$

Using the recursive structure in (G.15), the sequence $\left(\frac{\partial y_{0,\tau}}{\partial y_{s,2}}\right)_{s\in\mathbb{A}}$ of elements in that summation is a τ-times iterated convolution of the policy sequences $c_{T-2}, \ldots, c_{T-\tau+1}$ where

$$\frac{\partial y_{0,\tau}}{\partial y_{s,2}} \quad = \quad \sum_{b_1} \cdots \sum_{b_{t-1}} c_{b_1, T-(\tau-1)} \cdots c_{s-(b_1+\cdots+b_{t-1}), T-2} \tag{G.17}$$

and hence is an absolutely convergent and monotonically decreasing (on both sides of zero) sequence (see also the iterative derivation in (A.7) in the proof of Theorem 1). Moreover, the same convolution argument we used in (G.10) and (G.11) applied to the sequences in (G.15) and (G.16) yields that the rate at which they converge to zero is an increasing function of $|b|$. Consequently, since the monotonicity argument holds for each element in the summand, it also holds for the discounted sum in (G.14), i.e., for any $b \neq 0$

$$r^{\gamma}_{|b|+1} \;=\; \left(\frac{\gamma_{|b|+1,T}}{\gamma_{|b|,T}}\right) \;\leq\; \left(\frac{\gamma_{|b|+2,T}}{\gamma_{|b|,T}}\right) \;=\; r^{\gamma}_{|b|+2}$$

establishing the weak monotonicity of the cross-sectional rates for γ_T.

Given this property, to show that c_T satisfies the strict monotonicity of the cross-sectional rates, one uses the exact same deconstruction we used in equations (G.10) and (G.11) followed by monotone likelihood arguments. So, we leave it to the reader. Therefore, this concludes the proof of the statement in the first part of Lemma 5.

Proof of the second part. Let $G_m \subset G$ be as defined in the first part, namely, $G_m := \{(c_a) \in G \mid \frac{c_{|a|+1}}{c_{|a|}}$ is weakly increasing in $|a|\}$. We know from the first part that the right hand side of (G.18) is a self-map on G_m and that it maps any $(c_{a,T}) \in G_m$ to $(c'_{a,T}(\alpha_3)) \in G_m$ as defined

$$c'_{a,T}(\alpha_3) \quad := \quad \Delta_T^{-1}\left[\alpha_1\, I_{\{a\in\{-1,1\}\}} + \sum_{b\neq 0}\gamma_{b,T}(\alpha_3)\, c_{a-b,T}\right] \tag{G.18}$$

We are making the dependence on α_3 explicit to be able to do the comparative statics exercise across different values for α_3. Our objective is to show that given $\alpha_3' > \alpha_3$, $r_{a,T}(\alpha_3') > r_{a,T}(\alpha_3)$, for any $a \neq 0$. To accomplish that, it will suffice to show that the right hand side of (G.18) maps any element of G_m to a sequence with higher convergence rates uniformly across agents under α_3' than under α_3. Since the unique equilibrium sequence lies in G_m, this would show that the equilibrium sequence would possess the monotonicity property relative to α_3 as well. The proof is once again by induction on the number of periods T.

For $T = 2$, we know from (G.9) that

$$r^{\gamma}_{|b|+1,2} = \left(\frac{\gamma_{|b|+1,T}}{\gamma_{|b|,T}}\right) = r_1$$

and from Theorem 2-(ii) (see also footnote 19) that r_1 is strictly increasing in α_3, where $r_1 = \left(\frac{1}{2\,\alpha_3}\right) - \sqrt{\left(\frac{1}{2\,\alpha_3}\right)^2 - 1}$. Using once again the deconstruction in (G.10), for any $(c_{a,2})_{a\in\mathbb{A}} \in G_m$

$$\begin{aligned}
\Delta_2\, c'_{a+1,2} \quad &= \quad \left[(\gamma_{1,2}c_{a-1,2})\, r^{\gamma}_{2,2} + (\gamma_{2,2}c_{a-2,2})\, r^{\gamma}_{3,2} + \ldots\right] \\
&\quad +[(\gamma_{1,2}c_{a-1,2})\, r_{a,2} + (\gamma_{-1,2}c_{a+1,2})\, r_{a+2,2} + (\gamma_{-2,2}c_{a+2,2})\, r_{a+3,2} + \ldots] \\
&= \quad r_{a+1,2}\sum_{b\neq 0}\gamma_{b,2}\, c_{a-b,2} \\
&= \quad \Delta_2\, r'_{a+1,2}\, c'_{a,2}
\end{aligned} \tag{G.19}$$

$r'_{a+1,2}$ is strictly increasing in α_3 since $r^{\gamma}_{b+1,2}$ is strictly increasing in α_3, for any $b \geq 2$, inside the first brackets in the first line, and $r^{\gamma}_{|b|,2}$ is increasing in $|b|$ as we showed in the first part. So, in equilibrium, $r_{|a|+1,2} = \frac{c_{|a|+1,2}}{c_{|a|,2}}$ is necessarily strictly increasing in α_3 as well.

Now, assume for induction that $(\gamma_l)_{l=2}^{T-1}$ and $(c_l)_{l=2}^{T-1}$ are strictly increasing in α_3. Using the same convolution argument in (G.15) and (G.16), the convergence rate for γ_T would be strictly increasing in α_3 since all future period convergence rates $(c_l)_{l=2}^{T-1}$ are strictly increasing in α_3. Namely, for $\alpha_3' > \alpha_3$,

$$r^{\gamma}_{b,T}(\alpha_3') \quad > \quad r^{\gamma}_{b,T}(\alpha_3), \qquad \text{for any } b \neq 0. \tag{G.20}$$

Using one last time the deconstruction we used in (G.19) in equilibrium for T

$$\begin{aligned}
\Delta_T\, c_{a+1,T} \quad &= \quad \left[(\gamma_{1,T}c_{a-1,T})\, r^{\gamma}_{2,T} + (\gamma_{2,T}c_{a-2,T})\, r^{\gamma}_{3,T} + \ldots\right] \\
&\quad +[(\gamma_{1,T}c_{a-1,T})\, r_{a,T} + (\gamma_{-1,T}c_{a+1,T})\, r_{a+2,T} + (\gamma_{-2,T}c_{a+2,T})\, r_{a+3,T} + \ldots] \\
&= \quad r_{a+1,T}\sum_{b\neq 0}\gamma_{b,T}\, c_{a-b,T} \\
&= \quad \Delta_T\, r_{a+1,T}\, c_{a,T}
\end{aligned}$$

$r'_{a+1,T}$ is strictly increasing in α_3 since $r^{\gamma}_{b+1,T}$ is strictly increasing in α_3, for any $b \geq 2$, inside the first brackets in the first line, and $r^{\gamma}_{|b|,T}$ is increasing in $|b|$ as we showed in the first part. So, in equilibrium, $r_{|a|+1,T} = \frac{c_{|a|+1,T}}{c_{|a|,T}}$ is necessarily strictly increasing in α_3 as well. This concludes the proof. ■

Proof of Lemma 6: Our the environment being symmetric, it suffices to study the first order condition of a single agent, agent zero, in a $T-$period problem

$$\begin{aligned} 0 &= \alpha_1 (y_{0,0} - y_{0,1}) + \alpha_2 (\theta_{0,1} - y_{0,1}) + \alpha_3 (y_{-1,1} - y_{0,1}) + \alpha_3 (y_{1,1} - y_{0,1}) \\ &\quad + E\left[\sum_{\tau=2}^{T} \beta^{\tau-1} \left(-\alpha_1 \, (y_{0,\tau-1} - y_{0,\tau}) \frac{\partial}{\partial y_{0,1}} (y_{0,\tau-1} - y_{0,\tau}) + \alpha_2 \, (\theta_{0,\tau} - y_{0,\tau}) \frac{\partial}{\partial y_{0,1}} y_{0,\tau} \right.\right. \\ &\quad \left.\left. - \alpha_3 \, (y_{-1,\tau} - y_{0,\tau}) \frac{\partial}{\partial y_{0,1}} (y_{-1,\tau} - y_{0,\tau}) - \alpha_3 \, (y_{1,\tau} - y_{0,\tau}) \frac{\partial}{\partial y_{0,1}} (y_{1,\tau} - y_{0,\tau}) \right) \middle| \left(y^{t-1}, \theta^t\right) \right] \end{aligned} \tag{G.21}$$

For $T = 1$, using the first order condition and substituting for the policy coefficients multiplying $y_{a,0}$, we obtain

$$\alpha_1 \left(I_{\{a=0\}} - c_{a,1} \right) - \alpha_2 \, c_{a,1} + \alpha_3 \, (c_{a+1,1} + c_{a-1,1} - 2c_{a,1}) = 0$$

Summing both sides over a yields

$$\alpha_1 - (\alpha_1 + \alpha_2) \, C_1 = 0$$

hence

$$C_1 = \frac{\alpha_1}{\alpha_1 + \alpha_2}$$

Since the structural equations are the same, by using the same arguments, we also get $D_1 = \frac{\alpha_2}{\alpha_1+\alpha_2}$.

For $T = 2$, using the first-order condition and substituting for the policy coefficients multiplying $y_{a,0}$, we obtain once again

$$\begin{aligned} 0 &= \alpha_1 \, (\mathbf{1}\{a = 0\} - c_{a,2}) - \alpha_2 \, c_{a,2} + \alpha_3 \, (c_{a+1,2} + c_{a-1,2} - 2c_{a,2}) \\ &\quad + \beta \Bigg[- \alpha_1 (1 - c_{0,1}) (c_{a,2} - \sum_{b_1} c_{b_1,1} \, c_{a-b_1,2}) - \alpha_2 \, c_{0,1} \sum_{b_1} c_{b_1,1} \, c_{a-b_1,2} \\ &\quad - \alpha_3 (c_{-1,1} - c_{0,1}) \sum_{b_1} (c_{b_1+1,1} - c_{b_1,1}) \, c_{a-b_1,2} - \alpha_3 (c_{1,1} - c_{0,1}) \sum_{b_1} (c_{b_1-1,1} - c_{b_1,1}) \, c_{a-b_1,2} \Bigg] \end{aligned}$$

Since all sequences are absolutely summable, summing over a and using the fact that $C_1 = \frac{\alpha_1}{\alpha_1+\alpha_2}$

$$\begin{aligned} 0 &= \alpha_1 \, (1 - C_2) - \alpha_2 \, C_2 + C_2 \beta \Big[- \alpha_1 (1 - c_{0,1}) (1 - C_1) - \alpha_2 \, c_{0,1} \, C_1 \Big] \\ &= \alpha_1 - C_2 \, (\alpha_1 + \alpha_2) + \beta \, C_2 \Big[- \alpha_1 (1 - C_1) + c_{0,1} \, (\alpha_1 - C_1 \, (\alpha_1 + \alpha_2)) \Big] \\ &= \alpha_1 - C_2 \, (\alpha_1 + \alpha_2) - \beta \, C_2 \, \alpha_1 \, (1 - C_1) \end{aligned}$$

Hence, solving for C_2 yields

$$C_2 = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1 \beta (1 - C_1)} < \frac{\alpha_1}{\alpha_1 + \alpha_2} = C_1$$

Assume now that the result is true up to $T-1$. To show that it holds for T as well, we use the first-order condition and substitute for the policy coefficients multiplying $y_{a,0}$. Now, consider only the last period ($t=T$) terms inside the brackets in the summation in (G.21), namely

$$\beta^{T-1}\Bigg[-\alpha_1 \frac{\partial}{\partial y_{0,1}}(y_{0,T-1}-y_{0,T})\times$$

$$\left(\sum_{b_2}\cdots\sum_{b_{T-2}} c_{b_2,2}\cdots c_{b_{T-2},T-1}\, c_{a-(b_2+\ldots+b_{T-2}),T} - \sum_{b_1}\cdots\sum_{b_{T-1}} c_{b_1,1}\cdots c_{b_{T-1},T-1}\, c_{a-(b_1+\ldots+b_{T-1}),T}\right)$$

$$-\alpha_2 \frac{\partial y_{0,T}}{\partial y_{0,1}} \sum_{b_1}\cdots\sum_{b_{T-1}} c_{b_1,1}\cdots c_{b_{T-1},T-1}\, c_{a-(b_1+\ldots+b_{T-1}),T}$$

$$-\alpha_3 \frac{\partial}{\partial y_{0,1}}(y_{-1,T}-y_{0,T})\times$$

$$\left(\sum_{b_1}\cdots\sum_{b_{T-1}} c_{b_1,1}\cdots c_{b_{T-1},T-1}\, c_{a+1-(b_1+\ldots+b_{T-1}),T} - \sum_{b_1}\cdots\sum_{b_{T-1}} c_{b_1,1}\cdots c_{b_{T-1},T-1}\, c_{a-(b_1+\ldots+b_{T-1}),T}\right)$$

$$-\alpha_3 \frac{\partial}{\partial y_{0,1}}(y_{1,T}-y_{0,T})\times$$

$$\left(\sum_{b_1}\cdots\sum_{b_{T-1}} c_{b_1,1}\cdots c_{b_{T-1},T-1}\, c_{a-1-(b_1+\ldots+b_{T-1}),T} - \sum_{b_1}\cdots\sum_{b_{T-1}} c_{b_1,1}\cdots c_{b_{T-1},T-1}\, c_{a-(b_1+\ldots+b_{T-1}),T}\right)\Bigg]$$

and summing over a yields

$$\begin{aligned}
& \beta^{T-1}\left[-\alpha_1 \frac{\partial}{\partial y_{0,1}}(y_{0,T-1}-y_{0,T})\,(C_2\cdots C_T - C_1C_2\cdots C_T) - \alpha_2 \frac{\partial y_{0,T}}{\partial y_{0,1}} C_1\cdots C_T\right] \\
= \quad & C_2\cdots C_T\,\beta^{T-1}\left[-\alpha_1 \frac{\partial}{\partial y_{0,1}}(y_{0,T-1}-y_{0,T})\,(1-C_1) - \alpha_2 \frac{\partial y_{0,T}}{\partial y_{0,1}} C_1\right] \\
= \quad & C_2\cdots C_T\,\beta^{T-1}\left[-\alpha_1 \frac{\partial y_{0,T-1}}{\partial y_{0,1}}\,(1-C_1) + \frac{\partial y_{0,T}}{\partial y_{0,1}}\,(\alpha_1(1-C_1) - \alpha_2\, C_1)\right] \\
= \quad & C_2\cdots C_T\,\beta^{T-1}\left[-\alpha_1 \frac{\partial y_{0,T-1}}{\partial y_{0,1}}\,(1-C_1)\right] \qquad \text{(G.22)}
\end{aligned}$$

where we use the fact that $C_1 = \frac{\alpha_1}{\alpha_1+\alpha_2}$ in the third line. The sum of terms multiplied by α_3 cancel out.

Applying the same method to the period $T-1$ term and adding the period T term in (G.22) to it yields

$$\beta^{T-2}\left[-\alpha_1 \frac{\partial}{\partial y_{0,1}} (y_{0,T-2} - y_{0,T-1}) \, (C_3 \cdots C_T - C_2 C_3 \cdots C_T) - \alpha_2 \frac{\partial y_{0,T-1}}{\partial y_{0,1}} C_2 \cdots C_T\right]$$
$$+ C_2 \cdots C_T \, \beta^{T-1}\left[-\alpha_1 \frac{\partial y_{0,T-1}}{\partial y_{0,1}} \, (1-C_1)\right]$$
$$= C_3 \cdots C_T \, \beta^{T-2}\left[-\alpha_1 \frac{\partial}{\partial y_{0,1}} (y_{0,T-2} - y_{0,T-1}) \, (1-C_2) - \alpha_2 \frac{\partial y_{0,T-1}}{\partial y_{0,1}} C_2 - \alpha_1 \beta \frac{\partial y_{0,T-1}}{\partial y_{0,1}} \, (1-C_1) \, C_2\right]$$
$$= C_3 \cdots C_T \, \beta^{T-2}\left[-\alpha_1 \frac{\partial y_{0,T-2}}{\partial y_{0,1}} \, (1-C_2) + \frac{\partial y_{0,T-1}}{\partial y_{0,1}} (\alpha_1 (1-C_2) \, - \alpha_2 C_2 - \alpha_1 \beta \, (1-C_1) \, C_2)\right]$$
$$= C_3 \cdots C_T \, \beta^{T-2}\left[-\alpha_1 \frac{\partial y_{0,T-2}}{\partial y_{0,1}} \, (1-C_2) + \frac{\partial y_{0,T-1}}{\partial y_{0,1}} (\alpha_1 - C_2 \, (\alpha_1 + \alpha_2 + \alpha_1 \beta (1-C_1)))\right]$$
$$= C_3 \cdots C_T \, \beta^{T-2}\left[-\alpha_1 \frac{\partial y_{0,T-2}}{\partial y_{0,1}} \, (1-C_2)\right]$$

where, in the line before the last one, we used the fact that $C_2 = \frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_1\beta(1-C_1)}$. One can use induction by iterating this argument across period $T-2, \ldots, 2$ brackets and apply recursively the hypothesis that $C_\tau = \frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_1\beta(1-C_{\tau-1})}$ for $\tau < T$. This way, the first-order condition in (G.21) can be rolled back to

$$\begin{aligned} 0 &= \alpha_1(1-C_T) - \alpha_2 C_T \\ &\quad + C_T \, \beta \left[-\alpha_1 \left(1 - \frac{\partial y_{0,2}}{\partial y_{0,1}}\right)(1-C_{T-1}) - \alpha_2 \frac{\partial y_{0,2}}{\partial y_{0,1}} C_{T-1} - \alpha_1 \beta \frac{\partial y_{0,2}}{\partial y_{0,1}} C_{T-1}(1-C_{T-2})\right] \\ &= \alpha_1(1-C_T) - \alpha_2 C_T \\ &\quad + C_T \, \beta \left[-\alpha_1(1-C_{T-1}) + \frac{\partial y_{0,2}}{\partial y_{0,1}} \left[\alpha_1(1-C_{T-1}) - \alpha_2 C_{T-1} - \alpha_1 \beta C_{T-1}(1-C_{T-2})\right]\right] \end{aligned}$$

and applying once again the hypothesis that, for $T-1$, $C_{T-1} = \frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_1\beta(1-C_{T-2})}$, we obtain

$$0 = \alpha_1(1-C_T) - \alpha_2 \, C_T + C_T \, \beta \left[-\alpha_1(1-C_{T-1})\right]$$

and solving for C_T gives

$$C_T = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1 \beta (1-C_{T-1})}$$

Moreover, since by hypothesis

$$C_1 > \ldots > C_{T-1}$$

we also get

$$C_T = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1 \beta (1-C_{T-1})} < \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1 \beta (1-C_{T-2})} = C_{T-1}$$

as claimed in the statement. Finally, $(C_T)_{T=1}^{\infty}$ formed by the above construction is a monotone decreasing, bounded sequence of real numbers. Therefore, by the Monotone Convergence Theorem (see e.g. Bartle (1976), p.105), $C_T \to C_\infty$, which solves

$$C_\infty = \frac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_1 \beta (1-C_\infty)}$$

The arguments are identical for D_T and D_∞. This concludes the proof. ■

Proof of Lemma 7: -*Uniqueness of the solution to the system (G.1).* We will first argue that knowing the values of $c = (c_b)_{b\in\mathbb{A}}$, there exists a unique solution to the system in (G.1). Let $\mathbb{A}_k$ represent the finite truncation of the set of agents which includes agent 0 and his k closest neighbors on the left and on the right. For $T = \infty$, we rewrite the system in (G.1) as: for any $a \in \mathbb{A}_k$ and with $\gamma_0 := \Delta_\infty$,

$$-\alpha_1 I_{\{a=0\}} \;=\; -\gamma_0 c_a + \sum_{b\in\mathbb{A}_k\setminus\{0\}} \gamma_b c_{a-b}$$

which can be written equivalently as

$$I_{\{a=0\}} \;=\; -\gamma_0 \alpha_1^{-1} c_a + \sum_{b\in\mathbb{A}_k\setminus\{0\}} \gamma_b \alpha_1^{-1} c_{a-b} \tag{G.23}$$

under the assumption that $\alpha_1 \neq 0$, as assumed in Theorem 3.

Consider now the circulant matrix[52] M_k whose first row is given by the transpose of the column vector $c = (c_b)_{b\in\mathbb{A}_k}$ and each of its rows is a once right-shifted version of the previous one, i.e. $M_k(i+1,j) = M_k(i,j-1)$.[53] With this definition, the system in (G.23) can be written as

$$e \;=\; -\alpha_1^{-1} M_k \gamma_k \tag{G.24}$$

where $e := (1,0,\ldots,0)$ and the vector γ_k is such that $\gamma_{kb} = \gamma_b$ for $b \in \mathbb{A}_k \setminus \{0\}$ and $\gamma_{k0} = -\gamma_0$. We know from Lemma 1.1 (vi) in Carmona et al. (2015) that a circulant matrix M_k is invertible if e is in its column space. As equation (G.24) shows, e is indeed in the column space of M_k and the unique vector of weights $-\alpha_1^{-1}\gamma_k$ combining the columns of M_k generates it. This shows that, for any $k > 1$, M_k identifies the vector $-\alpha_1^{-1}\gamma_k$ associated with the policy coefficient sequence $c = (c_b)_{b\in\mathbb{A}_k}$. Next, we show that this continues to hold for the whole set of agents $\mathbb{A}$ as well.

The policy coefficient sequences $c = (c_b)_{b\in\mathbb{A}}$ are positive and absolutely convergent sequences. So, the infinite circulant matrix M is bounded as $\sum_j |M(i,j)| = \sum_i |M(i,j)| = C < 1$, and $M\gamma$ is convergent for any $\gamma \in [0,1]^{\mathbb{A}}$. Hence, the sequence $-\alpha_1^{-1}\gamma$ is the unique solution of (G.24) with $\mathbb{A}_k = \mathbb{A}$, thanks to the result 3.2 in Cooke (1950). This shows that for any $k \in \{1,2,\ldots\} \cup \{\infty\}$, the sequence $c = (c_b)_{b\in\mathbb{A}_k}$ identifies the sequence $-\alpha_1^{-1}\gamma$ in the sense that the latter is the unique solution of (G.24).

-γ_b *is increasing in* β *for* $b \neq 0$. Let $\beta' > \beta$. Using its definition in (A.6), γ_0 can be written as

$$\begin{aligned}\gamma_0 \;=\; \Delta_\infty \;\;:=\;\; & \alpha_1 + \alpha_2 + 2\alpha_3 + \sum_{t=2}^{\infty} \beta^{t-1} \left(\alpha_1 \left(\frac{\partial}{\partial y_{a,1}} (y_{a,t-1} - y_{a,t}) \right)^2 + \alpha_2 \left(\frac{\partial}{\partial y_{a,1}} y_{a,t} \right)^2 \right. \\ & \left. + \;\; \alpha_3 \left(\frac{\partial}{\partial y_{a,1}} (y_{a-1,t} - y_{a,t}) \right)^2 + \alpha_3 \left(\frac{\partial}{\partial y_{a,1}} (y_{a+1,t} - y_{a,t}) \right)^2 \right) \;>\; 0\end{aligned} \tag{G.25}$$

[52]A **circulant matrix** is a special kind of Toeplitz matrix fully specified by one vector, which appears as one of the rows of the matrix. Each other row vector of the matrix is shifted one element to the right relative to the preceding row vector. See e.g. Davis (1970) for an in-depth discussion of circulant matrices.

[53]The operations of addition and subtraction are legitimate for $\mathbb{A}_k$ represented by a circle, defined modularly.

Since for any $t > 1$ and for $a \in \mathbb{A}$

$$\frac{\partial y_{a,t}}{\partial y_{a+b,1}} = \sum_{b_1} \cdots \sum_{b_{t-1}} c_{b_1} \cdots c_{b-(b_1+\cdots+b_{t-1})}, \tag{G.26}$$

all terms after β^{t-1} are constant positive terms that are functions of the policy coefficients. Hence, increasing β would increase the sum. Moreover, from the first part of the lemma, we know we can identify the sequence $(\alpha_1^{-1}\gamma_b)_{b\in\mathbb{A}}$. So, the relative ratios of the terms across agents should be constant, i.e., observational equivalence requires that

$$\frac{\gamma_b(\beta)}{\gamma_0(\beta)} = \frac{\gamma_b(\beta')}{\gamma_0(\beta')}$$

Hence, γ_b should be increasing in β, for any $b \neq 0$.

- $\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1}$ *is continuously differentiable and* $\frac{\partial}{\partial p}\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1} > 0$ *and* $\frac{\partial}{\partial \alpha_1}\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1} < 0$*, for any* $(\alpha_1, p, \beta(p))$. Using the definition in (A.9), we can write for $b \neq 0$

$$\begin{aligned}
\frac{\gamma_b(\alpha_1,\, p,\, \beta(p))}{\alpha_1} &= \left(\frac{p-\alpha_1}{2\alpha_1 p}\right) I_{\{b=-1,1\}} \\
&\quad + \sum_{t=2}^{\infty} \beta^{t-1} \Bigg[\left(\frac{\partial y_{a,t-1}}{\partial y_{a,1}} - \frac{\partial y_{a,t}}{\partial y_{a,1}}\right)\left(\frac{\partial y_{a,t}}{\partial y_{a+b,1}} - \frac{\partial y_{a,t-1}}{\partial y_{a+b,1}}\right) \\
&\quad - \left(\frac{1-p}{p}\right) \frac{\partial y_{a,t}}{\partial y_{a,1}} \frac{\partial y_{a,t}}{\partial y_{a+b,1}} \\
&\quad + \left(\frac{p-\alpha_1}{2\alpha_1 p}\right)\left(\frac{\partial y_{a,t}}{\partial y_{a,1}} - \frac{\partial y_{a-1,t}}{\partial y_{a,1}}\right)\left(\frac{\partial y_{a-1,t}}{\partial y_{a+b,1}} + \frac{\partial y_{a+1,t}}{\partial y_{a+b,1}} - 2\frac{\partial y_{a,t}}{\partial y_{a+b,1}}\right)\Bigg]
\end{aligned} \tag{G.27}$$

where, given that $p := \frac{\alpha_1}{\alpha_1+\alpha_2}$, we substitute $\frac{\alpha_2}{\alpha_1} = (1-p)\frac{(\alpha_1+\alpha_2)}{\alpha_1} = \left(\frac{1-p}{p}\right)$; we also use the normalization $\alpha_1 + \alpha_2 + 2\alpha_3 = 1$ to write $\frac{\alpha_3}{\alpha_1} = \frac{1}{2}(1-(\alpha_1+\alpha_2)) = \left(\frac{p-\alpha_1}{2\alpha_1 p}\right)$; finally we also use the symmetry of the policy coefficients around zero for the last line. We know that the expression in the last parenthesis in the last line is positive since

$$\begin{aligned}
\frac{\partial}{\partial y_{a+b,1}}\left(y_{a-1,t} + y_{a+1,t} - 2y_{a,t}\right) &= \frac{\partial}{\partial y_{a+b,1}} \sum_{b\in\mathbb{A}} c_b \left(y_{a-1+b,t-1} + y_{a+1+b,t-1} - 2y_{a+b,t-1}\right) \\
&= \frac{\partial}{\partial y_{a+b,1}} \sum_{b\in\mathbb{A}} \left[c_{b+1}\, y_{a+b,t-1} + c_{b-1}\, y_{a+b,t-1} - 2c_b\, y_{a+b,t-1}\right] \\
&= \sum_{b\in\mathbb{A}} \left[c_{b+1} + c_{b-1} - 2c_b\right] \frac{\partial y_{a+b,t-1}}{\partial y_{a+b,1}} \\
&> 0.
\end{aligned}$$

The first line is by applying the policy function; second line is by a change of variable thanks to the symmetry of the policy function across agents; the inequality is thanks to the convexity property of the policy coefficients in (ii) in (A.3), which in turn makes the expression inside the brackets positive.

The partial derivative of $\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1}$ with respect to p is well-defined and continuous since $\alpha_1 \neq 0$ (hence $p \neq 0$)

$$
\begin{aligned}
\frac{\partial}{\partial p}\frac{\gamma_b(\alpha_1,\, p,\, \beta(p))}{\alpha_1} &= \left(\frac{1}{2p^2}\right) I_{\{b=-1,1\}} \\
&+ \sum_{t=2}^{\infty} \beta^{t-1} \left[\left(\frac{1}{p^2}\right) \frac{\partial y_{a,t}}{\partial y_{a,1}} \frac{\partial y_{a,t}}{\partial y_{a+b,1}} \right. \\
&\left. + \left(\frac{1}{2p^2}\right) \left(\frac{\partial y_{a,t}}{\partial y_{a,1}} - \frac{\partial y_{a-1,t}}{\partial y_{a,1}}\right) \left(\frac{\partial y_{a-1,t}}{\partial y_{a+b,1}} + \frac{\partial y_{a+1,t}}{\partial y_{a+b,1}} - 2\frac{\partial y_{a,t}}{\partial y_{a+b,1}}\right) \right] \\
&+ \left(\frac{\partial \beta}{\partial p}\right) \sum_{t=2}^{\infty} (t-1)\beta^{t-2} \left[\left(\frac{\partial y_{a,t-1}}{\partial y_{a,1}} - \frac{\partial y_{a,t}}{\partial y_{a,1}}\right) \left(\frac{\partial y_{a,t}}{\partial y_{a+b,1}} - \frac{\partial y_{a,t-1}}{\partial y_{a+b,1}}\right) \right. \\
&- \left(\frac{1-p}{p}\right) \frac{\partial y_{a,t}}{\partial y_{a,1}} \frac{\partial y_{a,t}}{\partial y_{a+b,1}} \\
&\left. + \left(\frac{p-\alpha_1}{2\alpha_1 p}\right) \left(\frac{\partial y_{a,t}}{\partial y_{a,1}} - \frac{\partial y_{a-1,t}}{\partial y_{a,1}}\right) \left(\frac{\partial y_{a-1,t}}{\partial y_{a+b,1}} + \frac{\partial y_{a+1,t}}{\partial y_{a+b,1}} - 2\frac{\partial y_{a,t}}{\partial y_{a+b,1}}\right) \right] \\
&> 0
\end{aligned}
\tag{G.28}
$$

The first three lines yield the direct effect of the differential change in p whose sum is positive. The last three lines yield the indirect effect of the differential change in p through β using the chain rule and the expression $\beta(p \,|\, C_\infty) = \frac{p - C_\infty}{p\, C_\infty (1 - C_\infty)}$ from (C.5), whose sum is positive as well. This is because $\frac{\partial \beta}{\partial p} = \frac{1}{p^2(1-C_\infty)} > 0$; and the sum of all other expressions in the last three lines is the partial derivative of $\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1}$ with respect to β; this latter is positive as we proved above that $\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1}$ is increasing in β. Similarly, the partial derivative of $\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1}$ with respect to α_1 is well-defined

$$
\begin{aligned}
\frac{\partial}{\partial \alpha_1}\frac{\gamma_b(\alpha_1,\, p,\, \beta(p))}{\alpha_1} &= \left(-\frac{1}{2\alpha_1^2}\right) I_{\{b=-1,1\}} \\
&+ \sum_{t=2}^{\infty} \beta^{t-1} \left(-\frac{1}{2\alpha_1^2}\right) \left(\frac{\partial y_{a,t}}{\partial y_{a,1}} - \frac{\partial y_{a-1,t}}{\partial y_{a,1}}\right) \left(\frac{\partial y_{a-1,t}}{\partial y_{a+b,1}} + \frac{\partial y_{a+1,t}}{\partial y_{a+b,1}} - 2\frac{\partial y_{a,t}}{\partial y_{a+b,1}}\right) \\
&< 0
\end{aligned}
\tag{G.29}
$$

continuous, and strictly negative at any $(\alpha_1, p, \beta(p))$.

So far, we have proved that the partial derivatives are well-defined and continuous everywhere in the admissible domain. Therefore, the map $\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1}$ is continuously differentiable (see e.g., Theorem 41.2 on Class C^1 on p. 376 in Bartle (1976)). We have also shown that the the partial derivatives satisfy $\frac{\partial}{\partial p}\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1} > 0$ and $\frac{\partial}{\partial \alpha_1}\frac{\gamma_b(\alpha_1, p, \beta(p))}{\alpha_1} < 0$, for any $(\alpha_1, p, \beta(p))$. This concludes the proof of Lemma 7(i).

-*There exist agents $b \neq b'$ such that for any $(\alpha_1, p, \beta(p))$,*

$$
-\left.\frac{\frac{\partial \alpha_1^{-1}\gamma_{b'}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b'}}{\partial \alpha_1}}\right|_{\alpha_1^{-1}\gamma_{b'}(\alpha_1,p,\beta(p))=\bar{\gamma}_b'} > -\left.\frac{\frac{\partial \alpha_1^{-1}\gamma_{b}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b}}{\partial \alpha_1}}\right|_{\alpha_1^{-1}\gamma_{b}(\alpha_1,p,\beta(p))=\bar{\gamma}_b} > 0. \tag{G.30}
$$

This result studies more closely the marginal rate of substitution (MRS) between α_1 and p that maintain the same levels for the maps

$$\frac{\gamma_b(\alpha_1,\, p,\, \beta(p))}{\alpha_1} - \bar{\gamma}_b \quad = \quad 0, \qquad \text{for } b \in \mathbb{A}. \tag{G.31}$$

Thanks to Lemma 7(i), the Implicit Function Theorem (see e.g., Bartle (1976), p.384) states that, for any $b_1 \in \mathbb{A}$, there exists a continuous and differentiable function $\alpha_1(p\,|\,\mathcal{P}_{b_1})$ that gives the unique value of α_1 for any value of p such that $(\alpha_1, p, \beta(p)) \in \mathcal{P}_{b_1}$, the level set defined in (C.6). The slope of $\alpha_1(p\,|\,\mathcal{P}_{b_1})$ is the marginal rate of substitution between α_1 and p that sustain the level in (G.31). Since by Lemma 7(i), $\frac{\partial}{\partial\alpha_1}\frac{\gamma_b(\alpha_1,\, p,\, \beta(p))}{\alpha_1} < 0$, the total derivative of the implicit function gives

$$\frac{\partial \alpha_1^{-1}\gamma_b}{\partial \alpha_1}\, d\alpha_1 + \frac{\partial \alpha_1^{-1}\gamma_b}{\partial p}\, dp \quad = \quad 0$$

or equivalently, the MRS between α_1 and p on the same level curve yields

$$MRS^b_{\alpha_1,p} \;=\; \frac{d\alpha_1}{dp} \;=\; -\left.\frac{\frac{\partial \alpha_1^{-1}\gamma_{b_1}}{\partial p}}{\frac{\partial \alpha_1^{-1}\gamma_{b_1}}{\partial \alpha_1}}\right|_{\mathcal{P}_{b_1}} \;>\; 0. \tag{G.32}$$

is positive thanks to Lemma 7(i), which we proved above. We have an overdetermined system with a total of $|\mathbb{A}|$ restrictions thanks to the implicit functions in (G.31) and two parameters (α_1, p) to determine. In particular, we will be interested in its behavior across agents as $|b|$ gets large. For $b \neq 0$, the partial derivative $\frac{\partial}{\partial\alpha_1}\frac{\gamma_b(\alpha_1,\, p,\, \beta(p))}{\alpha_1}$ we computed in (G.29) is

$$\begin{aligned}\frac{\partial}{\partial \alpha_1}\frac{\gamma_b(\alpha_1,\, p,\, \beta(p))}{\alpha_1} \;=\; & -\frac{1}{2\alpha_1^2}\Big[I_{\{b=-1,1\}} \\ & + \sum_{t=2}^{\infty}\beta^{t-1}\left(\frac{\partial y_{a,t}}{\partial y_{a,1}} - \frac{\partial y_{a-1,t}}{\partial y_{a,1}}\right)\left(\frac{\partial y_{a-1,t}}{\partial y_{a+b,1}} + \frac{\partial y_{a+1,t}}{\partial y_{a+b,1}} - 2\frac{\partial y_{a,t}}{\partial y_{a+b,1}}\right)\Big]\end{aligned} \tag{G.33}$$

and the partial derivative $\frac{\partial}{\partial\alpha_1}\frac{\gamma_b(\alpha_1,\, p,\, \beta(p))}{\alpha_1}$ we computed in (G.28) is

$$\begin{aligned}\frac{\partial}{\partial p}\frac{\gamma_b(\alpha_1,\, p,\, \beta(p))}{\alpha_1} \;=\; & \frac{1}{p^2}\Bigg(\left(\frac{1}{2}\right)I_{\{b=-1,1\}} + \sum_{t=2}^{\infty}\beta^{t-1}\Bigg[\frac{\partial y_{a,t}}{\partial y_{a,1}}\,\frac{\partial y_{a,t}}{\partial y_{a+b,1}} \\ & + \left(\frac{1}{2}\right)\left(\frac{\partial y_{a,t}}{\partial y_{a,1}} - \frac{\partial y_{a-1,t}}{\partial y_{a,1}}\right)\left(\frac{\partial y_{a-1,t}}{\partial y_{a+b,1}} + \frac{\partial y_{a+1,t}}{\partial y_{a+b,1}} - 2\frac{\partial y_{a,t}}{\partial y_{a+b,1}}\right)\Bigg]\Bigg) \\ & + \left(\frac{\partial\beta}{\partial p}\right)\sum_{t=2}^{\infty}(t-1)\beta^{t-2}\Bigg[\left(\frac{\partial y_{a,t-1}}{\partial y_{a,1}} - \frac{\partial y_{a,t}}{\partial y_{a,1}}\right)\left(\frac{\partial y_{a,t}}{\partial y_{a+b,1}} - \frac{\partial y_{a,t-1}}{\partial y_{a+b,1}}\right) \\ & - \left(\frac{1-p}{p}\right)\frac{\partial y_{a,t}}{\partial y_{a,1}}\,\frac{\partial y_{a,t}}{\partial y_{a+b,1}} \\ & + \left(\frac{p-\alpha_1}{2\alpha_1 p}\right)\left(\frac{\partial y_{a,t}}{\partial y_{a,1}} - \frac{\partial y_{a-1,t}}{\partial y_{a,1}}\right)\left(\frac{\partial y_{a-1,t}}{\partial y_{a+b,1}} + \frac{\partial y_{a+1,t}}{\partial y_{a+b,1}} - 2\frac{\partial y_{a,t}}{\partial y_{a+b,1}}\right)\Bigg]\end{aligned}$$

Since for any $t > 1$ and for $a \in \mathbb{A}$

$$\frac{\partial y_{a,t}}{\partial y_{a+b,1}} \quad = \quad \sum_{b_1}\cdots\sum_{b_{t-1}} c_{b_1}\cdots c_{b-(b_1+\cdots+b_{t-1})}, \tag{G.34}$$

in all three expressions, the terms inside the brackets involving $\frac{\partial y_{a,t-1}}{\partial y_{a,1}}$, $\frac{\partial y_{a,t}}{\partial y_{a,1}}$, and $\frac{\partial y_{a-1,t}}{\partial y_{a,1}}$ are iterated convolution of the policy sequence $(c_b)_{b\in\mathbb{A}}$, basically constant coefficients that do not change as b changes. The terms that change are the ones involving $\frac{\partial y_{a,t-1}}{\partial y_{a+b,1}}$, $\frac{\partial y_{a,t}}{\partial y_{a+b,1}}$, $\frac{\partial y_{a-1,t}}{\partial y_{a+b,1}}$, and $\frac{\partial y_{a+1,t}}{\partial y_{a+b,1}}$, as $|b|$ gets larger.

We know from Lemma 5 that the rate at which the policy coefficients converge to zero forms a convergent sequence that is monotonically increasing in $|b|$. Moreover, the same convolution argument we used in (G.10) and (G.11) in the proof of Lemma 5 can be used for the expression in (G.34) to show that the rate at which they converge to zero forms a convergent sequence that is monotonically increasing in $|b|$. Hence, thanks to symmetry of the policy function, letting $r^y(b,t) := \frac{\frac{\partial y_{a,t}}{\partial y_{a+b,1}}}{\frac{\partial y_{a,t}}{\partial y_{a+b-1,1}}}$, one can write the terms inside the second parenthesis in (G.33) as

$$\frac{\partial y_{a-1,t}}{\partial y_{a+b,1}} + \frac{\partial y_{a+1,t}}{\partial y_{a+b,1}} - 2\frac{\partial y_{a,t}}{\partial y_{a+b,1}} \quad = \quad \frac{\partial y_{a,t}}{\partial y_{a+b,1}}\left(r^y(b+1,t) + \frac{1}{r^y(b-1,t)} - 2\right) \tag{G.35}$$

to show how their convergence rates are related to that of $\frac{\partial y_{a,t}}{\partial y_{a+b,1}}$. We will give the argument for $b > 0$ wlog since the one for $b < 0$ is identical. The expression inside the parenthesis is positive but decreasing in b. So, for any $b > 0$

$$\frac{\frac{\partial y_{a,t}}{\partial y_{a+b+1,1}}\left(r^y(b+2,t) + \frac{1}{r^y(b,t)} - 2\right)}{\frac{\partial y_{a,t}}{\partial y_{a+b,1}}\left(r^y(b+1,t) + \frac{1}{r^y(b-1,t)} - 2\right)} \quad < \quad \frac{\frac{\partial y_{a,t}}{\partial y_{a+b+1,1}}}{\frac{\partial y_{a,t}}{\partial y_{a+b,1}}} \quad = \quad r^y(b+1,t)$$

which means that the rate of convergence for two consecutive such terms is less than $r^y(b+1,t)$, the convergence rate of $\frac{\partial y_{a,t}}{\partial y_{a+b,1}}$ terms.

The denominator in the definition of $MRS^b_{\alpha_1,p}$ in (G.32) consists only of the discounted sum of such terms in (G.35) for all periods $t \geq 2$, which decline at a rate lower than $r^y(b+1,t)$. The numerator also incorporates these terms as well as others that converge at rates higher than $r^y(b+1,t)$. Based on these facts, applying to the discounted sums above, tedious but straightforward modifications of the monotone convergence arguments we used in (G.10) and (G.11) in the proof of Lemma 5 yields that the overall convergence rate of the numerator is higher than the overall convergence rate of the denominator at any given $b \neq 0$. This in turn implies that the $MRS^b_{\alpha_1,p}$ is increasing in $|b|$, which means that the rate at which p is substituted for α_1 to make the levels in (G.31) intact can be ranked across agents b. Therefore, there exist b and b' for which the statement in Lemma 7(ii) is true. This concludes the proof. ■

H Social Welfare

We provide here the formal welfare arguments for finite economies with i.i.d. preference shocks for the clarity of the intuition delivered. The extension of the line of proof to more general processes and to the infinite-horizon is tedious but straightforward. We first write the planning problem recursively. For any agent $a \in \mathbb{A}$, for all $t = 1, \ldots, T$, and all $(y_{T-1}, \theta^t) \in \mathbf{Y} \times \mathbf{\Theta}^t$, let the value of using the choice rule h in the continuation be defined as

$$V^{h,T-(t-1)}\left(R^a\, y_{T-1}, R^a\, \theta_t\right) = -\alpha_1 \left(y_{a,t-1} - h_{T-(t-1)}(R^a\, y_{t-1}, R^a\, \theta_t)\right)^2$$
$$-\alpha_2 \left(\theta_{a,t} - h_{T-(t-1)}(R^a\, y_{t-1}, R^a\, \theta_t)\right)^2$$
$$-\alpha_3 \left(h_{T-(t-1)}(R^{a-1}\, y_{t-1}, R^{a-1}\, \theta_t) - h_{T-(t-1)}(R^a\, y_{t-1}, R^a\, \theta_t)\right)^2$$
$$-\alpha_3 \left(h_{T-(t-1)}(R^{a+1}\, y_{t-1}, R^{a+1}\, \theta_t) - h_{T-(t-1)}(R^a\, y_{t-1}, R^a\, \theta_t)\right)^2$$
$$+\beta \int V^{h,T-t} \left(R^a\, \left\{h_t(R^b\, y_{t-1}, R^b\, \theta_t)\right\}_{b\in\mathbb{A}}, R^a\, \theta_{t+1}\right) \mathbb{P}\left(d\theta_{t+1}\right)$$

which leads us to the following definition

Definition 4 (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 $g : \mathbf{Y} \times \mathbf{\Theta} \times \{1, \ldots, T\} \to Y$ is said to be **efficient** if it is a solution, for all $a \in \mathbb{A}$, and for all $t = 1, \ldots, T$, to*

$$\operatorname*{arg\,max}_{\{h\in CB(\mathbf{Y}\times\mathbf{\Theta},Y)\}} \int \Bigg[- \alpha_1 \left(y_{0,t-1} - h_{T-(t-1)}(R^a\, y_{t-1}, R^a\, \theta_t)\right)^2$$
$$-\alpha_2 \left(\theta_{a,t} - h_{T-(t-1)}(R^a\, y_{t-1}, R^a\, \theta_t)\right)^2$$
$$-\alpha_3 \left(h_{T-(t-1)}(R^{a-1}\, y_{t-1}, R^{a-1}\, \theta_t) - h_{T-(t-1)}(R^a\, y_{t-1}, R^a\, \theta_t)\right)^2$$
$$-\alpha_3 \left(h_{T-(t-1)}(R^{a+1}\, y_{t-1}, R^{a+1}\, \theta_t) - h_{T-(t-1)}(R^a\, y_{t-1}, R^a\, \theta_t)\right)^2$$
$$+\beta\, V^{h,T-t} \left(R^a\, \left\{h(R^b\, y_{t-1}, R^b\, \theta_t)\right\}_{b\in\mathbb{A}}, R^a\, \theta_{t+1}\right) \Bigg] \mathbb{P}\left(d\theta_t\right) \mathbb{P}\left(d\theta_{t+1}\right) \pi_t\left(dy_{t-1}\right)$$

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

We will use continuity arguments so endow the underlying space $\mathbf{Y} \times \mathbf{\Theta}$ with the product topology. Product topology is metrizable, say by metric d[54]. In the final period of this finite horizon economy, with absolutely continuous distribution π_{T-1} on the space of choice profiles y_{T-1}[55] with a positive density, the planner maximizes ex-ante (before the realization of θ_T) the expected utility of a given agent, say of agent $0 \in \mathbb{A}$, by choosing a symmetric policy function $h \in CB(\mathbf{Y} \times \mathbf{\Theta}, Y)$, the space of bounded, continuous, and Y-valued measurable functions. [56] The space $\mathbf{Y} \times \mathbf{\Theta}$ is compact with respect to the product topology since Y 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.

[54]Let $|\cdot|$ be the usual Euclidean norm. For any $(y, \theta), (y', \theta') \in \mathbf{Y} \times \mathbf{\Theta}$, let

$$d\left((y, \theta), (y', \theta')\right) := \sum_{a\in\mathbb{A}} 2^{-a} \left(|y_a - y'_a| + |\theta_a - \theta'_a|\right)$$

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

[55]Starting with an initial π_0 which is absolutely continuous, the MPE policy function and the absolutely continuous preference shocks induce a sequence (π_t) of absolutely continuous distributions on t-period equilibrium choice profiles.

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

Lemma 8 *For any* $(y_{T-1}, \theta_T) \in \mathbf{Y} \times \mathbf{\Theta}$,

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

Proof: The proof uses an extension of the usual calculus of variation techniques to our symmetric strategic environment. We prove it for the class of bounded, continuous, and measurable, real-valued functions on $\mathbf{Y} \times \mathbf{\Theta}$. Then, we use the restriction of the result to a subset of it, the space of bounded, continuous, and measurable, Y-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.,

$$\begin{aligned} J(a) &:= \int u\left(y_{0,T-1}, (h+ak)(y_{T-1}, \theta_T), (h+ak)(R^{-1}\, y_{T-1}, R^{-1}\,\theta_T),\right. \\ &\quad \left.(h+ak)(R\, y_{T-1}, R\,\theta_T), \theta_{0,T}\right) \mathbb{P}\left(d\theta_T\right) \pi_{T-1}\left(dy_{T-1}\right) \end{aligned}$$

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

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

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

$$\begin{aligned} J'(0) &:= \int \Big[u_2 \left(y_{0,T-1}, h(y_{T-1}, \theta_T), h(R^{-1}\, y_{T-1}, R^{-1}\,\theta_T), h(R\, y_{T-1}, R\,\theta_T), \theta_{0,T}\right)\, k(y_{T-1}, \theta_T) \\ &+u_3 \left(y_{0,T-1}, h(y_{T-1}, \theta_T), h(R^{-1}\, y_{T-1}, R^{-1}\,\theta_T), h(R\, y_{T-1}, R\,\theta_T), \theta_{0,T}\right)\, k(R^{-1}y_{T-1}, R^{-1}\theta_T) \\ &+u_4 \left(y_{0,T-1}, h(y_{T-1}, \theta_T), h(R^{-1}\, y_{T-1}, R^{-1}\,\theta_T), h(R\, y_{T-1}, R\,\theta_T), \theta_{0,T}\right)\, k(Ry_{T-1}, R\theta_T) \Big] \\ &\times \mathbb{P}\left(d\theta_T\right) \pi_{T-1}\left(dy_{T-1}\right) \;=\; 0 \end{aligned}$$

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

$$\begin{aligned} 0 \;\neq\; & u_2 \left(\hat{y}_0, h(\hat{y}, \hat{\theta}), h(R^{-1}\,\hat{y}, R^{-1}\,\hat{\theta}), h(R\,\hat{y}, R\,\hat{\theta}), \hat{\theta}_0\right) \\ & +u_3 \left(\hat{y}_1, h(R\,\hat{y}, R\,\hat{\theta}), h(\hat{y}, \hat{\theta}), h(R^2\,\hat{y}, R^2\,\hat{\theta}), \hat{\theta}_1\right) \\ & +u_4 \left(\hat{y}^{-1}, h(R^{-1}\,\hat{y}, R^{-1}\,\hat{\theta}), h(R^{-2}\,\hat{y}, R^{-2}\,\hat{\theta}), h(\hat{y}, \hat{\theta}), \hat{\theta}_{-1}\right) \end{aligned} \tag{H.1}$$

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 neigborhood $A \subset \mathbf{Y} \times \mathbf{\Theta}$ around $(\hat{y}, \hat{\theta})$ such that the above expression stays positive for all $(y_{T-1}, \theta_T) \in A$.[57] Assume that $a_1 = (\hat{y}, \hat{\theta})$, $a_2 = (R\,\hat{y}, R\,\hat{\theta})$, and $a_3 = (R^{-1}\,\hat{y}, R^{-1}\,\hat{\theta})$ are distinct points. Otherwise, since the underlying space Y 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

$$k(y,\theta) = k(R\,y, R\,\theta) = k(R^{-1}\,y, R^{-1}\,\theta) = \begin{cases} \gamma\,[\delta - d((y,\theta), a_1)]\,, & \text{if } (y,\theta) \in A_1 \\ 0, & \text{otherwise.} \end{cases} \tag{H.2}$$

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 function[58]. Substitute k into equation (H.1). 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 $\mathbf{Y} \times \mathbf{\Theta}$ for which the expression (H.1) 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. ■

This implies that

$$\begin{aligned} h(y_{T-1}, \theta_T) \quad &= \quad (\alpha_1 + \alpha_2 + 4\alpha_3)^{-1} \Big(\alpha_1\, y_{0,T-1} + \alpha_2\, \theta_{0,T-1} \\ &\quad + 2\alpha_3\, h(R^{-1}\, y_{T-1}, R^{-1}\,\theta_T) + 2\alpha_3\, h(R\, y_{T-1}, R\,\theta_T)\Big) \end{aligned} \tag{H.3}$$

As in the proof of existence, the operator induced by (H.3) is a contraction on the Banach space of bounded, continuous, measurable functions with the supnorm, whose unique fixed point is in G, defined in (A.3). Therefore, one can fit the following solution

$$h(y_{T-1}, \theta_T) = \sum_a c_a^P\, y_{a,T-1} + \sum_a d_a^P\, \theta_{a,T}$$

substituting, we get

$$\begin{aligned} \sum_a c_a^P\, y_{a,T-1} + \sum_a d_a^P\, \theta_{a,T} \quad &= \quad (\alpha_1 + \alpha_2 + 4\alpha_3)^{-1} \Bigg[\alpha_1\, y_{0,T-1} + \alpha_2\, \theta_{0,T} \\ + 2\alpha_3 \left(\sum_a c_a^P\, y_{a-1,T-1} + \sum_a d_a^P\, \theta_{a-1,T}\right) &+ 2\alpha_3 \left(\sum_a c_a^P\, y_{a+1,T-1} + \sum_a d_a^P\, \theta_{a+1,T}\right)\Bigg] \end{aligned}$$

By matching coefficients, we get for all $a \in \mathbb{A}$

$$c_a^P = (\alpha_1 + \alpha_2 + 4\alpha_3)^{-1} \left[2\alpha_3\, c_{a-1}^P + 2\alpha_3\, c_{a+1}^P + \alpha_1 \mathbf{1}_{\{a=0\}}\right]$$

$$d_a^P = (\alpha_1 + \alpha_2 + 4\alpha_3)^{-1} \left[2\alpha_3\, d_{a-1}^P + 2\alpha_3\, d_{a+1}^P + \alpha_2 \mathbf{1}_{\{a=0\}}\right]$$

[57] Endowed with the product topology, the space $\mathbf{Y} \times \mathbf{\Theta}$ is metrizable by the metric d. See footnote 54. Product topology and the associated metric allows us to choose positive measure proper subsets of Y for choices of nearby agents and the whole sets Y and Θ for far away agents, staying at the same time in the close vicinity of the point $(\hat{y}, \hat{\theta})$.

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

The same method as in the proof of Theorem 2 yields for any $a \in \mathbb{A}$,

$$c_a^P = r_P^{|a|} \left(\frac{\alpha_1}{\alpha_1+\alpha_2}\right)\left(\frac{1-r}{1+r}\right) \quad \text{and} \quad d_a^P = r_P^{|a|} \left(\frac{\alpha_2}{\alpha_1+\alpha_2}\right)\left(\frac{1-r}{1+r}\right) \tag{H.4}$$

$$r_P = \left(\frac{\Delta_P}{4\,\alpha_3}\right) - \sqrt{\left(\frac{\Delta_P}{4\,\alpha_3}\right)^2 - 1} \quad \text{with} \quad \Delta_P = \alpha_1 + \alpha_2 + 4\,\alpha_3. \tag{H.5}$$

We next compare the equilibrium policy sequence in Theorem 2 (see also footnote 19) with the planner's optimal choice coefficient sequence. Notice that

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

which implies that $r_P > r_1$ since r_P is decreasing in $\frac{\Delta_P}{4\alpha_3}$ by (H.5). Thus, the planner's optimal policy coefficient sequence converges to zero slower than the equilibrium policy coefficient sequence. Moreover, the equilibrium policy cannot satisfy the FOC of the planner's problem. Therefore, the equilibrium is inefficient for finite-horizon economies.

I Extensions

General Neighborhood Network Structures: Section 2.3.1

The analogue for the expression (A.2) implied by the FOC in the final period of the economy with a general network structure is

$$y_{a,1} = \Delta_{a,1}^{-1} \left(\alpha_{a,1}\, y_{a,0} + \alpha_{a,2}\, \theta_{a,1} + \sum_{b \in N(a)} \alpha_{a,b}\, y_{b,1} \right) \tag{I.1}$$

where $\Delta_{a,1} := \alpha_{a,1} + \alpha_{a,2} + \sum_{b\in N(a)} \alpha_{a,b} > 0$. Hence, the best-response for agent a, $BR_1^a : G^{\mathbb{A}\setminus\{a\}} \to G$, is uniformly continuous for all $a \in \mathbb{A}$, which implies that so is the induced best-response profile $BR_1 = (BR_1^a) : G^{\mathbb{A}} \to G^{\mathbb{A}}$. The convex set $G^{\mathbb{A}}$ is also compact thanks to Lemma 3. Hence a fixed point exists (not necessarily unique). Thus, an equilibrium exists for a one-period economy. If the uniform boundedness condition (in footnote 31) holds, the rate at which each one of the maps in (I.1) contracts is uniformly bounded. Hence, BR_1 becomes a contraction operator on $G^{\mathbb{A}}$ implying uniqueness of the equilibrium for a one-period economy. For a $T < \infty$ period economy, one mimicks the induction arguments of the general existence proof to obtain the analogue of the expression (A.10) in Lemma 2, i.e.,

$$0 \quad = \quad -y_{a,1}\, \Delta_{a,T} + \alpha_{a,1}\, y_{a,0} + \alpha_{a,2}\, \theta_{a,1} + \sum_{b\neq a} \gamma_{a,b,T}\, y_{b,1} + \sum_{b\in\mathbb{A}} \sum_{\tau=2}^{T} \mu_{b,\tau,T}\, E\left[\theta_{a+b,\tau} | \theta^1\right] \tag{I.2}$$

where $\Delta_{a,T} := \alpha_{a,1} + \alpha_{a,2} + \sum_{b\neq a} \gamma_{a,b,T} + \sum_{b\in\mathbb{A}} \sum_{\tau=2}^{T} \mu_{b,\tau,T}$. Applying the above arguments to the system of equations (I.2) for each a gives existence of an equilibrium for a $T < \infty$ period economy. Since $\gamma_{a,b,T}$ is the total effect of a change in $y_{b,1}$ ($b \neq a$) on the expected discounted marginal utility of agent a (as in expression (A.9)), the sum of these effects is uniformly bounded across agents if the peer effects are uniformly bounded across agents. Hence, once again, the best-response profile for a T-period economy, BR_T, becomes a contraction on $G^{\mathbb{A}}$, implying uniqueness of equilibrium. Moreover, using the same arguments in the third step of the general existence proof in Appendix A, the sequence of unique equilibria approximates a stationary equilibrium as the horizon length becomes arbitrarily large. ■

Global Interactions in Section 2.3.3

The proof uses straightforward modifications of the arguments in Section 5 of Bisin, Horst, and Özgür (2006) to our environment. Interested reader should consult that work. ■

Social Accumulation of Habits in Section 2.3.2 a Similar to the arguments employed in General Neighborhood Network Structures, the expressions (A.2) and (A.10) change this time to

$$y_{a,1} = \Delta_{a,1}^{-1} \left(\alpha_{a,1}\, r_{a,1} + \alpha_{a,2}\, \theta_{a,1} + \sum_{b \in N(a)} \alpha_{a,b}\, y_{b,1} \right) \tag{I.3}$$

and

$$0 \quad = \quad -y_{a,1}\, \Delta_{a,T} + \alpha_{a,1}\, r_{a,1} + \alpha_{a,2}\, \theta_{a,1} + \sum_{b\neq a} \gamma_{a,b,T}\, y_{b,1} + \sum_{b\in\mathbb{A}} \sum_{\tau=2}^{T} \mu_{b,\tau,T}\, E\left[\theta_{a+b,\tau} | \theta^1\right] \tag{I.4}$$

respectively. The rest is a direct application of the arguments used above in the Section General Neighborhood Network Structures, because the coefficients in the new system of maps above are the same as those in (I.1) and (I.2). ■

J Details of Simulations

The core engine behind the simulations is a Matlab code, **g.m**, which computes the equilibrium policy weights recursively. The code is posted on Onur Özgür's webpage:

https://sites.google.com/site/onurozgurresearch/research,

and contains also detailed explanations. It uses as input parameters values of the preference parameters α_i, $i = 1, 2, 3$, the discount factor β, the horizon for the economy T, and the number of agents m on each side of a given agent so the total number of agents is $|\mathbb{A}| = 2m + 1$.

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

J.1 Details of the Monte Carlo Experiment

For this section, we assume that finite number of agents $\mathbb{A} = \{1, \ldots, N\}$ are placed on a circle. As seen in (9), for each $a \in \mathbb{A}$, the first-order condition of agent a's optimization problem admits the following simple expression:

$$y_{a,T} = \alpha_1 y_{a,T-1} + \alpha_3(y_{a-1,T} + y_{a+1,T}) + \alpha_2\gamma x_{a,T} + \varepsilon_{a,T} \tag{J.1}$$

where $\varepsilon_{a,T} = \alpha_2 u_{a,T}$. Since the correlation of regressors with the error term leads to inconsistency of least-squares methods, this equation can be consistently estimated, under assumptions 2 and 3, by using as instruments, for instance, own past observed shock $x_{a,T-1}$ and friends' current observed shocks $(x_{a-1,T} + x_{a+1,T})$. The 'exclusion restrictions' hold due to Assumption 2 and they yield the following population moment conditions, for any $a \in \mathbb{A}$

$$\begin{aligned} E\left[\varepsilon_{a,T} \mid x_{a,T-1}\right] &= 0 \\ E\left[\varepsilon_{a,T} \mid (x_{a-1,T} + x_{a+1,T})\right] &= 0 \end{aligned}$$

which can then be expressed simply, thanks to the Law of Iterated Expectations, as

$$\begin{aligned} E\left[x_{a,T-1}\,\varepsilon_{a,T}\right] &= 0 \\ E\left[(x_{a-1,T} + x_{a+1,T})\,\varepsilon_{a,T}\right] &= 0 \end{aligned}$$

This yields a total of $2N$ moment conditions for $t = T$ and 2 parameters, α_1, and α_3 to estimate. Similarly, the first-order condition at $T-1$, equation (10), is equivalent to the following econometric equation

$$\begin{aligned}[1+\beta\alpha_1(1-c_{1,1})]y_{a,T-1} \quad = \quad & \alpha_1 y_{a,T-2} + \alpha_3(y_{a-1,T-1} + y_{a+1,T-1}) + \beta(\alpha_1 - c_{1,1}(1-2\alpha_3))y_{a,T} \\ & + \gamma\alpha_2 x_{a,T-1} + \gamma\beta\alpha_2 c_{1,1} x_{a,T} + \varepsilon_{a,T-1} \end{aligned} \tag{J.2}$$

where $\varepsilon_{a,T-1}$ includes preference shocks as well as differences between expected and realized outcomes at T. Three valid instruments are enough to provide consistent estimates. When $\alpha_3 \neq 0$, these instruments could be, for instance, $x_{a,T-2}$, $(x_{a-1,T-1}+x_{a+1,T-1})$, and $(x_{a-2,T-1}+x_{a+2,T-1})$. Formally, the population moment conditions are, for any $a \in \mathbb{A}$

$$\begin{aligned} E\left[\varepsilon_{a,T-1} \mid x_{a,T-2}\right] &= 0 \\ E\left[\varepsilon_{a,T-1} \mid (x_{a-1,T-1} + x_{a+1,T-1})\right] &= 0, \\ E\left[\varepsilon_{a,T-1} \mid (x_{a-2,T-1} + x_{a+2,T-1})\right] &= 0 \end{aligned}$$

which can be written, for any $a, b \in \mathbb{A}$, thanks to the Law of Iterated Expectations, unconditionally as

$$\begin{aligned} & E\left[x_{a,T-2}\, \varepsilon_{a,T-1}\right] = 0 \\ & E\left[(x_{a-1,T-1} + x_{a+1,T-1})\, \varepsilon_{a,T-1}\right] = 0 \\ & E\left[(x_{a-2,T-1} + x_{a+2,T-1})\, \varepsilon_{a,T-1}\right] = 0 \end{aligned}$$

If we use all such possible combinations, we have $3N$ moment conditions for $t = T-1$ and 3 parameters, α_1, α_3, and β to estimate.

Given these $5N$ population moment conditions, nonlinear in the 3 parameters of interest, $\lambda := (\alpha_1, \alpha_3, \beta)$, the model is *overidentified*, and it is not possible to solve the system of analogous sample moment conditions for a unique value of the parameter vector. Instead, the structural coefficients can be estimated using the Generalized Method of Moments (GMM)[59].

Define the adjacency matrix $\mathbf{G}$ by

$$\mathbf{G}_{ij} = \begin{cases} 1, & \text{if } j = i-1, i+1 \\ 0, & \text{otherwise} \end{cases}$$

With $\mathbf{G}$ summarizing the overall interaction structure concisely, we can represent the instrument vectors as

- Friends' observable covariates $(x_{a-1,T} + x_{a+1,T})_{a\in\mathbf{A}}$ and $(x_{a-1,T-1} + x_{a+1,T-1})_{a\in\mathbf{A}}$, at time T and $T-1$, can be represented by $\mathbf{G}\mathbf{x}_T$ and $\mathbf{G}\mathbf{x}_{T-1}$, respectively.
- Friends's friends' observable covariates (excluding agent a himself) at time $T-1$, namely $(x_{a-2,T-1} + x_{a+2,T-1})_{a\in\mathbf{A}}$, can be represented by $\left(\mathbf{G}^2 - 2\mathbf{I}\right)\mathbf{x}_{T-1}$.

[59] We follow closely the setup of Section 6.5 of Cameron and Trivedi (2005) on nonlinear instrumental variables.

Finally, let's define the $5N \times 2N$ instruments matrix $\mathbf{Z}$ and the $2N \times 1$ error vector ε as

$$\mathbf{Z} \;:=\; \underbrace{\begin{bmatrix} \text{diag}(\mathbf{x}_{T-1}) & \mathbf{0} \\ \text{diag}(\mathbf{G}\mathbf{x}_T) & \mathbf{0} \\ \mathbf{0} & \text{diag}(\mathbf{x}_{T-2}) \\ \mathbf{0} & \text{diag}(\mathbf{G}\mathbf{x}_{T-1}) \\ \mathbf{0} & \text{diag}((\mathbf{G}^2 - 2\mathbf{I})\mathbf{x}_{T-1}) \end{bmatrix}}_{5N\times 2N} \quad \text{and} \quad \varepsilon := \underbrace{\begin{pmatrix} \varepsilon_T \\ \varepsilon_{T-1} \end{pmatrix}}_{2N\times 1} \tag{J.3}$$

where diag $(\mathbf{x})$ is the $N \times N$ diagonal matrix whose only non-zero entries on the diagonal are the elements of the vector $\mathbf{x}$, i.e.

$$\text{diag}\,(\mathbf{x}) \;:=\; \underbrace{\begin{bmatrix} \mathbf{x}_1 & & & & \\ & \ddots & & & \\ & & \mathbf{x}_a & & \\ & & & \ddots & \\ & & & & \mathbf{x}_N \end{bmatrix}}_{N\times N} \tag{J.4}$$

This way, we can write the $5N \times 1$ system of population moment conditions as

$$E\,[\,\mathbf{Z}\varepsilon\,] = \underbrace{\mathbf{0}}_{5N\times 1} \tag{J.5}$$

Using the analogy principle for finite samples, the corresponding sample moment can be written as

$$\frac{1}{M}\sum_{m=1}^{M} \mathbf{Z}_m \varepsilon_m = \mathbf{0} \tag{J.6}$$

where M is the sample size, i.e., the number of replica economies that we draw i.i.d. in each simple random sample. Since the number of moments is greater than the number of parameters, the model is overidentified and (J.6) has no solution for $\hat{\lambda}$. Instead, we choose $\hat{\lambda}$ so that a quadratic form in $M^{-1}\sum_{m=1}^{M} \mathbf{Z}_m \varepsilon_m$ is as close to zero as possible. Hence, the GMM estimator $\hat{\lambda}_{GMM}$ minimizes

$$Q_M(\lambda) \;=\; \left[\frac{1}{M}\sum_{m=1}^{M} \mathbf{Z}_m \varepsilon_m\right]' \mathbf{W}_M \left[\frac{1}{M}\sum_{m=1}^{M} \mathbf{Z}_m \varepsilon_m\right] \tag{J.7}$$

where the $5N \times 5N$ weighting matrix $\mathbf{W}_M$ is symmetric positive definite, and does not depend on λ. Solving for the optimal $\hat{\lambda}$ requires differentiating $Q_M(\lambda)$ in (J.7) with respect to λ to obtain the GMM first-order conditions

$$\left[\frac{1}{M}\sum_{m=1}^{M} \mathbf{Z}_m \left.\frac{\partial \varepsilon_m}{\partial \lambda}\right|_{\lambda=\hat{\lambda}}\right]' \mathbf{W}_M \left[\frac{1}{M}\sum_{m=1}^{M} \mathbf{Z}\varepsilon\right] \;=\; \underbrace{\mathbf{0}}_{3\times 1} \tag{J.8}$$

where we have multiplied by the scaling factor $1/2$ and where

$$\left.\frac{\partial \varepsilon}{\partial \lambda}\right|_{\lambda=\hat{\lambda}} \tag{J.9}$$

is the $2N \times 3$ Jacobian matrix for ε evaluated at the minimizer $\hat{\lambda}$, and where the Jacobian matrix of partial derivatives of $\varepsilon := (\varepsilon_T, \varepsilon_{T-1})$ with respect to the parameter vector $\lambda = (\alpha_1, \alpha_3, \beta)$ is

$$\frac{\partial \varepsilon}{\partial \lambda} = \begin{bmatrix} \frac{\partial \varepsilon_T}{\partial \alpha_1} & \frac{\partial \varepsilon_T}{\partial \alpha_3} & \frac{\partial \varepsilon_T}{\partial \beta} \\ \frac{\partial \varepsilon_{T-1}}{\partial \alpha_1} & \frac{\partial \varepsilon_{T-1}}{\partial \alpha_3} & \frac{\partial \varepsilon_{T-1}}{\partial \beta} \end{bmatrix} \tag{J.10}$$

such that

$$\begin{aligned} \frac{\partial \varepsilon_T}{\partial \alpha_1} &= -\mathbf{y}_{T-1} + \gamma\, \mathbf{x}_T \\ \frac{\partial \varepsilon_T}{\partial \alpha_3} &= -\mathbf{G}\mathbf{y}_T + 2\,\gamma\, \mathbf{x}_T \\ \frac{\partial \varepsilon_T}{\partial \beta} &= \mathbf{0} \end{aligned} \tag{J.11}$$

and

$$\begin{aligned} \frac{\partial \varepsilon_{T-1}}{\partial \alpha_1} &= \eta \left[\beta \left((1 - c_{11}) - \alpha_1 \frac{\partial c_{11}}{\partial \alpha_1} \right) (\mathbf{y}_{T-1} - \varepsilon_{T-1}) - \mathbf{y}_{T-2} - \beta \left(1 - (1 - 2\alpha_3) \frac{\partial c_{11}}{\partial \alpha_1} \right) \mathbf{y}_T \right. \\ &\qquad \left. + \gamma \mathbf{x}_{T-1} - \gamma\beta \left(\alpha_2 \frac{\partial c_{11}}{\partial \alpha_1} - c_{11} \right) \mathbf{x}_T \right] \\ \frac{\partial \varepsilon_{T-1}}{\partial \alpha_3} &= \eta \left[\alpha_1 \beta \frac{\partial c_{11}}{\partial \alpha_3} (\varepsilon_{T-1} - \mathbf{y}_{T-1}) - \mathbf{G}\mathbf{y}_{T-1} + \beta \left(\frac{\partial c_{11}}{\partial \alpha_3} (1 - 2\alpha_3) - 2c_{11} \right) \mathbf{y}_T + 2\gamma \mathbf{x}_{T-1} \right. \\ &\qquad \left. - \gamma\beta \left(\alpha_2 \frac{\partial c_{11}}{\partial \alpha_3} - 2c_{11} \right) \mathbf{x}_T \right] \\ \frac{\partial \varepsilon_{T-1}}{\partial \beta} &= \eta \left[\alpha_1 (1 - c_{11}) (\mathbf{y}_{T-1} - \varepsilon_{T-1}) - (\alpha_1 - c_{11} (1 - 2\alpha_3)) \mathbf{y}_T + \gamma \alpha_2 c_{11} \mathbf{x}_T \right] \end{aligned} \tag{J.12}$$

and $\alpha_2 = (1 - \alpha_1 - 2\alpha_3)$, $\eta := (1 + \beta\alpha_1(1 - c_{11}))^{-1}$, $r_1 = (2\alpha_3)^{-1} - \sqrt{(2\alpha_3)^{-2} - 1}$; $\frac{\partial c_{11}}{\partial \alpha_1} = \frac{r_1}{1-2\alpha_3} \frac{1-r_1}{1+r_1}$, and $\frac{\partial c_{11}}{\partial \alpha_3} = 2(1 - 2\alpha_3)^{-1} c_{11} + c_{11} \left(\frac{\partial r_1}{\partial \alpha_3} \right) \frac{1-2r_1-r_1^2}{r_1(1-r_1)^2}$ where $\left(\frac{\partial r_1}{\partial \alpha_3} \right) = \frac{1}{4\,a_3{}^3 \sqrt{\frac{1}{4\,a_3{}^2} - 1}} - \frac{1}{2\,a_3{}^2}$.

We know from Cameron and Trivedi (2005) Proposition 6.1 that the GMM estimator $\hat{\lambda}_{GMM}$, defined to be a root of the first-order conditions in (J.8), is consistent for λ_0, and is asymptotically normally distributed as $\sqrt{N}\left(\hat{\lambda}_{\mathbf{GMM}} - \lambda_{\mathbf{0}}\right) \xrightarrow{d} N(\mathbf{0}, \mathbf{V})$, a result due to Hansen (1982). We use a simple GMM estimator where the weight matrix $\mathbf{W}_M = \mathbf{I}_{5N}$. Since our replica economies are independent over m, we obtain the estimate $\widehat{\mathbf{S}}$ using as an obvious estimator

$$\widehat{\mathbf{S}} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{Z}_m \varepsilon_m (\mathbf{Z}_m \varepsilon_m)' \tag{J.13}$$

hence the GMM estimator $\hat{\lambda}_{GMM}$ is asymptotically normally distributed with mean λ_0 and with estimated asymptotic variance

$$\widehat{\mathbf{V}} = \left(\widehat{\mathbf{D}}' \widehat{\mathbf{S}}^{-1} \widehat{\mathbf{D}} \right)^{-1} \tag{J.14}$$

where

$$\widehat{\mathbf{D}} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{Z}_m \left. \frac{\partial \varepsilon_m}{\partial \lambda} \right|_{\lambda = \hat{\lambda}_{GMM}}. \tag{J.15}$$