Local Gossip and Intergenerational Family Transfers: Comparative Political Economy of Welfare Provision

Yuki Takagi*

March 8, 2011

Abstract

Why are generous transfers from the younger to the older generations made in some families and not in others. For example, people in northeast Japan tend to provide better nursing care to their aged parents and invest more in their children's education than those in the southwest. My paper argues that differences in intergenerational dependence are due to variation in community networks. My analysis of the sustainability of intergenerational transfers posits game theoretical models of overlapping generations in which breadwinners make transfers to their parents and children. A novel feature of my models is that there is a local community that may supply information about its members' past behaviors. I demonstrate that an efficient level of intergenerational transfers can be sustained if neighbors "gossip" about each other. As an implication, my theory suggests that individuals in a close-knit community will prefer lower levels of social protection. Empirical results from Japan support this argument: Individuals who interact with their neighbors tend to provide better nursing care to their aged parents, spend more on their childrens' education, and demand less from the government than those who do not interact with their neighbors.

*PhD candidate, Department of Government, Harvard University, Cambridge, MA 02138, e-mail: takagi@fas.harvard.edu.

1 Introduction

The level and means of welfare provision vary across societies and countries. Why are transfers from the younger to the older generations generous in some societies and not in others? Why does the family play a central role in care for the elderly in some countries while the government does so in others? For example, in southern European societies, the family functions as a welfare provider in terms of care for the elderly, childcare, and helping the unemployed (**???**). Moreover, extended families consisting of three or more generations have been no rarity in the region (Moreno 2006).

Furthermore, East Asian welfare regimes, which are often characterized by their low level of state-based transfers to the elderly and high levels of family-based transfers. To explain this East Asian variant of the welfare state, scholars of the region argue that Confucian ideology plays an important role (**??**). According to Confucian ethics, the family is an important source of welfare provision, with aged parents being cared for by children.

The Hong Kong special administrative region government, for example, has adopted Confucian ethics of filial piety to control social welfare costs (**?**). In fact, care for the elderly is primarily a family responsibility and the government is considered to be the last resort in Hong Kong. The family rather than the government has been regarded as the main source of welfare provision in Japan as well (Harada 1988). Thus, public welfare services play only a secondary role in the private welfare practice within the family (**?**). For example, public home care services are designated merely to support family care on the basis that the care for the elderly is their children's responsibility (Kono 2003).

The reliance on the welfare role of the family in East Asian countries contrasts sharply with the Western welfare system in which states play a central role. In addition to the cross-national differences, the level of family-based welfare provision varies even within the same country. For example, people in northeast Japan tend to provide better nursing care to their aged parents and invest more in their children's education than those in the southwest. Communities in the northeast tend to be small and isolated compared to those in the southwest.

Although it has been recognized that the family plays an important role in welfare provision, scholars of the field have focused on understanding the role of the state and market. As a result, little is known about the role and mechanisms of other institutions such as civic associations and the family. In this paper, I present a theoretical framework to analyze how family-based welfare provision is achieved and the role of the local community in the familial provision, with empirical findings supporting the argument. This theoretical framework may give a microfoundation to the "familialism" argument in the southern European welfare states and the "Confucian" argument in the East Asian welfare states.

In this paper, I argue that differences in the level of transfers within the family are due to demographic variation in community networks. To analyze the sustainability of intergenerational transfers, I propose game theoretical models of overlapping generations in which breadwinners make transfers to their parents and children. A novel feature of the models is that there is a local community that can supply information about its members' past behavior. I demonstrate that intergenerational transfers can be sustained if neighbors gossip about each other. My theory suggests that individuals in a close-knit community prefer lower levels of social protection. Empirical results from Japan support this argument: Individuals in small communities who interact with their neighbors tend to provide better nursing care to their aged parents, spend more on their children's education, and demand less from the government than those who do not interact with their neighbors.

This paper is organized as follows. In Section 2, I discuss the previous literature and the current paper's contribution to the literature. In Section 3, I present formal models and demonstrate that intergenerational transfers can be supported if neighbors gossip about each other. Empirical results in Section 4 support the argument: Individuals who interact with their neighbors tend to provide better nursing care to their aged parents, spend more on their children's education, and demand less from the government. In Section 5, I discuss this paper's implications and future extension.

2 Related Literature

2.1 Literature on the Welfare State

Scholars on the welfare state have placed emphasis on the state and the market, but the family has been largely discussed in the context of gender (Lewis 1992; O'Connor et al. 1999; Daly and Rake 2003; **?**). As a result, little is known about how *inter*-generational rather than *intra*-generational redistribution works within the family. This paper sheds new light on the role of the family in welfare provision.

Understanding the role of the family also helps us understand the East Asian variant of the welfare state, which is marked by a low level of social spending. Two approaches are used to explain the small welfare state in the region. The first approach explains the incentives of the governments. Wong (2004) ascribes the small social spending in East Asia to the absence of strong leftist parties and unions while Holiday (2000) argues that social spending in the region was minimized to promote economic growth. However, those authors fail to explain why the family works as a provider of welfare. They do not offer compelling accounts of the incentives to the family, such as why breadwinners are willing to provide care for their aged parents, for example.

The second approach, on the other hand, explains the incentives of family members. This group of scholars argues that heavy reliance on the family is possible because of Confucian values in East Asia (Jones 1993; Goodman et al. 1998). Their findings suggest the family is responsible for social protection rather than the state because of their values. This paper can be seen as falling within this second approach. While the previous literature simply assumes that people in the region have "Confucian" values and considers the values to be exogenous, this paper explains why the "Confucian" way of family transfers prevails in some regions and not in others. In this sense the present paper gives a microfoundation to the Confucian theory.

2.2 Literature on Social Capital

2.3 Theoretical Contributions

This present work is part of the literature on overlapping generations models (OLG models). Pioneered by **?**, OLG models have been widely studied by game theorists (**?????**)). **?**, for example, studies the ability of non-market institutions, such as the government and the family, to invest optimally in forward intergenerational goods (FIGs), such as education and environment, and backward intergenerational goods (BIGs), such as social security. BIGs transfer income from one generation to earlier generations; FIGs transfer income forward to younger generations. Most of the game theoretic OLG studies cited above assume that the entire history of the play of a game is common knowledge. That is, players are assumed to be informed of past events. In the context of families, it is assumed that each generation knows even its great-great-great-grandparent's past behavior, for example. Considering that the history may include actions taken by preceding generations before the current generations are born, this assumption of perfect information becomes questionable.

Very few attempts at relaxing this assumption have been made. **?**, however, does so and demonstrates that no intergenerational cooperation can be supported in pure-strategy equilibrium if information about the history of play is limited. The intuition is as follows: The limited information makes it impossible for each generation to condition its actions on the observed history in equilibrium, a condition crucial to support intergenerational cooperation in equilibrium. If older generations have better knowledge about past events than younger generations, the older generations can manipulate the information, an act that will be interpreted and acted on by younger generations by behaving as though different past events happened.

In response to **?**, **?** incorporate institutional features of overlapping generations organizations into Bhaskar's model to solve his impossibility theorem. Considering legislative bodies such as the U.S. Senate to be overlapping generations organizations, they show that the principals (the founding fathers) will agree to institute a mechanism that provides imperfectly informed legislators with the information about the history in the legislature, enabling intergenerational cooperation. Following **?**, the present paper provides another solution to Bhaskar's impossibility theorem in the setting of families. This paper demonstrates that intergenerational cooperation can be supported if neighbors gossip about each other. There are two key institutional differences between **?** and this present paper: First, this paper assumes that the local community stores the information as a summary statistic of information of unboundedly high order, while **?** assume that the institutionalized mechanism stores the information as the entire history of play in the legislature. Second, **?** endogenize the institutionalized mechanism, whereas the local community is exogenous in the current paper.

In addition to the literature on OLG models, this present work can be viewed as part of the literature on community enforcement, which is usually modeled as a repeated games with random matching. The literature on community enforcement can be divided into two strands based on the assumptions about players' knowledge. The first strand assumes that players know the past plays to some extent: **?** assumes that all players know the history of all matches in the population; **?**, Ellison (1994), and Harrington (1995) assume that players' knowledge is limited to the matches in which they have been directly involved; Rosenthal (1979), Landau (1979), Klein

(1992), Greif (1993, 1994), Tirole (1996), and Takahashi (2008) assume that players only have only first-order information; that is, they have information about their partners' past play but do not know their partners' past partner's past play.

The present work falls into a second strand, which assumes that each player is labeled with a status that is observable to his or her partner, and that a player can condition his or her action on the partner's status (Kandori 1992; Okuno-Fujiwara and Postlewaite 1995). A player's status is updated based on the realized action profile of the stage game and the player's and the partner's status in the previous period. Because a player's status at the next period depends on his or her current partner's status, which, in turn, depends on the partner's previous partner's status, and so on, a status is a summary statistic of information of unboundedly high order. The equilibrium in this setting is called a norm equilibrium. This paper introduces the concept of norm equilibrium to overlapping-generations games.

In terms of the role of the third party in sustaining cooperation, this paper is related to Milgrom et al. (1990), who investigate how a judge (a law merchant) serves to facilitate cooperation, while the current paper explores how a local community does so. Both this paper and that of Milgrom et al. may be viewed as attempts to investigate the role of the third party as an information device in sustaining cooperation.

3 Formal Models

This paper investigates (1) whether intergenerational transfers can be sustained when individuals are imperfectly informed about past events in their families and (2) whether and how a local community and its social norm contribute to the intergenerational family transfers. Throughout this paper, I assume all generations are selfish and are not altruistic.[1].

[1]I do not argue that individuals are completely selfish and no transfers within families derives from altruism. Individuals may make transfers simply because they care about family members. The amount of transfers which they make only with altruistic motives alone may not, however, be as high as the one they would make with both altruistic and selfish motives in cooperative equilibrium. In this paper, I focus on the transfers made out of selfish motives and do not address the those made for altruistic reasons.

3.1 Two-Generations Model: Backward Intergenerational Transfers from Children to Parents

In this subsection, I propose an overlapping generations model, using two generations to analyze backward intergenerational transfers, such as nursing care and financial support provided by individuals to their parents. I consider a situation in which a breadwinner decides how much to invest in a backward intergenerational good (BIG) that benefits only an elderly dependent. The key informational feature is that it is assumed that children can observe several preceding generations' behavior, such as how their parents treated their grandparents, but information about their distant ancestors' behavior, such as how their great-grandparents treated their great-great-grandparent, is unavailable. I first demonstrate that intergenerational transfers cannot be sustained under limited information and then show that close-knit communities serve to facilitate cooperation between generations.

3.1.1 Basic Model

Consider an infinitely-lived family with two generations alive in each period t ($t = 1, 2, \ldots$). At every period t, a single player, called generation t, is born to the family and lives for two periods: t and $t+1$. I call generation t the breadwinner in period t and the elderly dependent in period $t+1$. In period 1 there are a breadwinner 1 who stays in periods 1 and 2, and an elderly dependent 0 who stays only for that period.

The breadwinner has positive endowment, and the elderly dependent has endowment that is normalized to zero. Every period t, the breadwinner t decides how much to transfer to the elderly dependent $t-1$. Let $a_t \in A_t$ denote the amount transferred to the current elderly dependent, $t-1$, where action spaces are common across generations, $A_t = A$ for all t. The commodity is assumed to be infinitely divisible. The elderly dependent has no choices to make.

Generation t's utility function, $u : A_t \times A_{t+1} \to \mathbb{R}$, is decreasing in the transfer made in his middle-age and increasing in the transfer he receives in his old age. That is, a generation's lifetime utility depends only on the action taken while a breadwinner and the action that the breadwinner takes when he or she is old. Utility function $u(\cdot)$ satisfies the condition:

Assumption 1. $\forall \mathbf{a}, \mathbf{a}' \in A^2$, $u(\mathbf{a}) = u(\mathbf{a}') \Rightarrow \mathbf{a} = \mathbf{a}'$.

This condition ensures that $\operatorname*{argmax}_{a_t \in A_t} u(a_t, a_{t+1})$ is unique. Assuming that $\operatorname*{argmax}_{a_t \in A_t} u(a_t, a_{t+1})$ is independent of a_{t+1}, I label it $a_t = 0$, which can be interpreted as no transfers. Note that this overlapping generation game has a unique Markov equilibrium, in which every player chooses 0.

Let $h_t = (a_1, \dots, a_{t-1})$ denote the history of preceding actions taken until period t and $H_t (= A^{t-1})$ denote the set of all possible histories at t. I define the default informational environment as follows:

Definition 1. *For any $t = 1, 2, \dots$, generation t has* ***m-th order information*** *if he knows the actions of the last m generations, $(a_{t-1}, \dots, a_{t-m})$, but not any action taken prior to $t - m$, $(a_{t-m-1}, \dots, a_1)$.*

Assumption 2. *There exists a natural number m such that generation t has m-th order information for all $t = 1, 2, \dots$.*

If (h_t, h'_t) is any pair of histories that differ only in the actions taken by some of players i $(\leq t - m - 1)$, then the histories observed by generation t are identical for h_t and h'_t. For example, the first-order information is a record of the preceding generation's past play. That is, the information is limited in the sense that a generation does not know the actions taken in the family prior to its parents.

A pure strategy for generation t is a function $s_t : A_{t-m} \times \cdots \times A_{t-1} \to A_t$. Let S_t be the set of t's pure strategies. A strategy profile is an infinite sequence $(s_t)_{t=1}^{\infty}$ where $s_t \in S_t$ for all t. Thus, $s_t(a_{t-m}, \dots, a_{t-1})$ is the element of A_t which is induced by the observed history $(a_{t-m}, \dots, a_{t-1})$ when s_t is played.

An action by generation t, a_t, when he observes $(a_{t-m}, \dots, a_{t-1})$, induces an observed history for $t+1$, $(a_{t-m+1}, \dots, a_t)$. Different actions by generation t generate different information for $t+1$, and generation $t+1$ varies her actions with her observed history, $(a_{t-m+1}, \dots, a_t)$.

Adopting the terminology of **?**, I define our equilibrium definition as follows. Because past actions do not directly affect current or future utility, I do not have to deal with any beliefs regarding the histories.

Definition 2. *A strategy profile $(s_t)_{t=1}^{\infty}$ is a sequentially rational equilibrium if $\forall t$, $\forall (a_{t-m}, \dots, a_{t-1}) \in A_{t-m} \times \cdots \times A_{t-1}$, $\forall a_t \in A_t$, $u(s_t, s_{t+1} | a_{t-m}, \dots, a_{t-1}) \geq u(a_t, s_{t+1} | a_{t-m}, \dots, a_{t-1})$ where*

$$u(a_t, s_{t+1} | a_{t-m}, \dots, a_{t-1}) = u[a_t, s_{t+1}(a_{t-m+1}, \dots, a_t)]$$

and

$$u(s_t, s_{t+1}|a_{t-m}, \ldots, a_{t-1}) = u[s_t(a_{t-m}, \ldots, a_{t-1}), s_{t+1}(a_{t-m+1}, \ldots, a_t)].$$

The following theorem is a variant of Theorem 1 in **?**.

Theorem 1. *Under Assumption 1, the overlapping generations game with m-th order information has a unique pure strategy equilibrium where no intergenerational transfer is made.*

Proof. Suppose that generation t's equilibrium strategy s_t is based on action taken by generation $t-m$. Then, there exist $a_{t-m}, a'_{t-m} (a_{t-m} \neq a'_{t-m})$ such that $u_t(s_t, s_{t+1}|a_{t-m}, \ldots, a_{t-1}) > u_t(s_t, s_{t+1}|a'_{t-m}, \ldots, a_{t-1})$. Since s_{t+1} conditions only on actions taken by generations $t-m+1, \ldots, t$ and is independent of generation $t-m$'s action, generation t can improve his payoff by choosing $s_t(a_{t-m}, \ldots, a_{t-1})$ instead of $s_t(a'_{t-m}, \ldots, a_{t-1})$ when he observes $(a'_{t-m}, \ldots, a_{t-1})$. Hence, generation t's equilibrium strategy is not conditioned on $t-m$'s actions. Similarly, generation t's equilibrium strategy does not condition on k's action for any $k < t$. Thus, generation t's equilibrium strategy is not based on history. Therefore generation t's best response is $a_t = 0$, which forms a unique pure strategy equilibrium where no intergenerational transfer is made. □

Theorem 1 indicates that when individuals have limited information about past events, no strategy profile (which does not have to be Markov) in which players condition their behavior on the observed history, which is payoff-irrelevant, constitutes an equilibrium. Thus, intergenerational cooperation cannot be supported in pure-strategy equilibrium when information about past events is limited.

Intuitively, the limited information makes it impossible for each generation to condition its actions on the observed history, which is crucial to intergenerational cooperation. When the information is limited, individuals have better knowledge about past events than their children. Thus, by behaving as though different past events happened, they can manipulate the information that will be interpreted by their children. For instance, suppose that $m = 2$ and each generation plays the following strategy: Make transfers to the parent unless the parent abandoned the grandparent who made transfers in his middle-age. Suppose also that an individual knows that both his parent and grandparent did not make transfers. The above strategy tells him to take good care of his or her parent. However, knowing that his or her child will follow the above strategy, he or she can neglect the aged parent without being punished by the child in his or her old age. The key point is that he or she can deviate from the strategy profile without being caught by the child. More generally, any strategy profile that is conditioned on history fails to be an equilibrium, which prevents intergenerational cooperation.

3.1.2 Local Communities

In our basic model, it is shown that BIGs cannot be supported by pure-strategy equilibria when information is limited. In reality, however, we observe intergenerational family transfers. The question is whether they are all based on altruistic motives.

As a solution to this limited information problem, I modify the basic model by adding a local community that provides information about the history of play in each family through gossiping. In a close-knit community, it is natural to assume that intergenerational family transfers are observed by neighbors. Parents may complain about their children to neighbors, or neighbors may be able to tell the level of transfers from the well-being of the parents. As a result, neighbor gossip may provide the breadwinner with information necessary to see whether the parent's past behavior can be justified.

I consider a situation in which families in the previous basic model reside in a close-knit community. In a close-knit community, the breadwinner has a reputation based on how he or she treated his or her parent, and that reputation will become known to his or her child through his neighbors' gossip. The transfers the breadwinner will receive from his or her child in old age may depend on that his reputation, which subsequently determines the child's reputation. In this setting, BIGs can be supported by pure-strategy equilibria. Formally, this situation is modeled as follows.

Before introducing a local community to the basic model, I define the efficient level of transfers. Let $\operatorname*{argmax}_{a\in A} u(a,a) = 1$ be the efficient transfer, which I assume to be different from 0, the individually optimal action. Otherwise, it is trivial that no transfer is sustained in equilibrium. From Assumption 1,

Lemma 1. $u(0,1) > u(1,1) > u(0,0) > u(1,0)$

In each period t, the breadwinner t is assigned an element x_t of a finite set $X_t = \{\text{“good”}, \text{“bad”}\}$ which I refer to as the breadwinner's *status* or *status label*. A generation's status label $x_t \in X_t$ is determined through $\tau : X_{t-1} \times A_t \to X_t$. τ specifies the status label of generation t in the next period, $\tau(x_{t-1}, a_t) \in X_t$, when his previous generation's status label is $x_{t-1} \in X_{t-1}$ and his current action is $a_t \in A_t$. Because a generation's status depends on its parent's status, which in turn depends on the grandparent's status and so on, a status is a summary statistics of information of unboundedly high order. I call τ a *social norm* and a social norm is common knowledge.

A social norm τ is **family reciprocity** if

$$x_t = \tau(x_{t-1}, a_t) = \begin{cases} \text{“bad”} & \text{if } x_{t-1} = \text{“good” } \& \ a_t \neq 1 \\ \text{“good”} & \text{otherwise} \end{cases}$$

and $x_0 = $ “good”. Individuals are considered “bad” only when they do not take care of their parents whose status label is “good”.

Generation t is labeled by a social norm τ and the community informs the succeeding generation $t+1$ of t's status label x_t; that is, the status label of the elderly dependent is known to the breadwinner at period $t+1$. In particular, the breadwinner's action choice is typically a function of the previous generation's status label.

The history of intergenerational transfers may not be known; it becomes known to each generation only to the extent that it is reflected in the status labels of the elderly dependent.[2] A pure (Markov) strategy for a generation is a mapping $s_t : X_{t-1} \to A_t \ \forall t$ specifying a choice of action $s_t(x_{t-1})$ when the previous generation's status label is $x_{t-1} \in X_{t-1}$.

Our equilibrium definition here slightly differs from the last one in the information possessed by the decision makers. Because past actions do not directly affect current or future utility, I do not have to deal with any beliefs regarding the histories.

Definition 3. *A strategy profile $(s_t)_{t=1}^{\infty}$ is a sequentially rational equilibrium if*

$$\forall t, \ \forall x_{t-1} \in X_{t-1}, \ \forall a \in A, \ \ u(s_t, s_{t+1} | x_{t-1}) \geq u(a, s_{t+1} | x_{t-1})$$

where

$$u(a_t, s_{t+1} | x_{t-1}) = u[a_t, s_{t+1}(\tau(a_t, x_{t-1}))]$$

[2]Individuals may not know labels of their grandparents or more distant ancestors.

and

$$u(s_t, s_{t+1}|x_{t-1}) = u[s_t(x_{t-1}), s_{t+1}(\tau(s_t(x_{t-1}), x_{t-1}))].$$

I say that a strategy profile $(s_t)_{t=1}^{\infty}$ is a **tit-for-tat strategy** if

$$s_t(x_{t-1}) = \begin{cases} 1 & \text{if } x_{t-1} = \text{“good”} \\ 0 & \text{if } x_{t-1} = \text{“bad”}. \end{cases}$$

Theorem 2. *In a close-knit community whose social norm is family reciprocity, an efficient level of intergenerational transfers can be sustained as a sequentially rational equilibrium by a tit-for-tat strategy.*

Proof. Each generation is labeled by a social norm τ, i.e., $x_t = \tau(x_{t-1}, a_t)$ for all t. Consider a community whose social norm is family reciprocity, that is, the breadwinner has to take care of his or her parent to be labeled “good” the his parent is “good”.

Suppose that each generation plays a tit-for-tat strategy. Take an arbitrary period t. When $x_{t-1} =$ “good”, generation t's utility is

$$\begin{aligned} u(s_t, s_{t+1}|x_{t-1}) &= u(1, 1) \\ u(a_t, s_{t+1}|x_{t-1}) &= u(a_t, 0) \quad \text{if } a_t \neq 1 \end{aligned}$$

where $u(1, 1) > u(a_t, 0)$ for all $a_t \neq 1$. When $x_{t-1} =$ “bad”, generation t's utility is

$$\begin{aligned} u(s_t, s_{t+1}|x_{t-1}) &= u(0, 1) \\ u(a_t, s_{t+1}|x_{t-1}) &= u(a_t, 1) \quad \text{for all } a_t \end{aligned}$$

where $u(0, 1) \geq u(a_t, 1)$. Hence,

$$\forall t, \ \forall x_{t-1} \in X_{t-1}, \ \forall a \in A \ \ u(s_t, s_{t+1}|x_{t-1}) \geq u(a, s_{t+1}|x_{t-1}).$$

□

Theorem 3 establishes that in a community whose social norm is family reciprocity, an efficient level of intergenerational transfers can be sustained as a sequentially rational equilibrium by a tit-for-tat strategy. This result is in a sharp contrast with the case without a local community (Theorem 1) where BIGs cannot be supported by pure-strategy equilibrium.

The intuition is as follows. In the basic model, intergenerational cooperation is not sustainable because of the informational advantage of the older generation over the younger generation (Theorem 1). A close-knit community enables intergenerational cooperation because it nullifies the informational advantage of the older generation. Because a status label is a summary statistic of information of unboundedly high orders, it fills the information gap between the older generation and the younger generation so that the older generation cannot manipulate. Thus, by incorporating local communities, this present work demonstrates that the conclusion of the impossibility of intergenerational cooperation under limited information does not persist.

By Theorems 1 and 3 , the intergenerational cooperation is sustained in a close-knit community. Local gossip by neighbors (or extended families) serve to facilitate cooperation between generations.

3.1.3 Local Communities 2

In our basic model, it is shown that BIGs cannot be supported by pure-strategy equilibria when information is limited. In reality, however, we observe intergenerational family transfers. The question is whether they are all based on altruistic motives.

As a solution to this limited information problem, I modify the basic model by adding a local community that provides information about the history of play in each family through gossiping. In a close-knit community, it is natural to assume that intergenerational family transfers are observed by neighbors. Parents may complain about their children to neighbors, or neighbors may be able to tell the level of transfers from the well-being of the parents. As a result, neighbor gossip may provide the breadwinner with information necessary to see whether the parent's past behavior can be justified.

I consider a community that consists of two families. In this community, the breadwinner of each family has a reputation based on how he or she treated his or her parent, and that reputation will become known to his or her child through his neighbors' gossip. The transfers the breadwinner will receive from his or her child in old age may depend on that his reputation, which subsequently determines the child's reputation. In this setting, BIGs can be supported by pure-strategy equilibria. Formally, this situation is modeled as follows.

Before introducing a local community to the basic model, I define the efficient level of transfers. Let $\operatorname*{argmax}_{a\in A} u(a,a) = 1$ be the efficient transfer, which

I assume to be different from 0, the individually optimal action. Otherwise, it is trivial that no transfer is sustained in equilibrium. From Assumption 1,

Lemma 2. $u(0,1) > u(1,1) > u(0,0) > u(1,0)$

In each period t, the breadwinner t is assigned an element x_t of a finite set $X_t = \{$"good", "bad"$\}$ which I refer to as the breadwinner's *status* or *status label*. A generation's status label $x_t \in X_t$ is determined through $\tau : X_{t-1} \times A_t \to X_t$. τ specifies the status label of generation t in the next period, $\tau(x_{t-1}, a_t) \in X_t$, when his previous generation's status label is $x_{t-1} \in X_{t-1}$ and his current action is $a_t \in A_t$. Because a generation's status depends on its parent's status, which in turn depends on the grandparent's status and so on, a status is a summary statistics of information of unboundedly high order. I call τ a *social norm* and a social norm is common knowledge.

A social norm τ is **family reciprocity** if

$$x_t = \tau(x_{t-1}, a_t) = \begin{cases} \text{“bad”} & \text{if } x_{t-1} = \text{“good” } \& \ a_t \neq 1 \\ \text{“good”} & \text{otherwise} \end{cases}$$

and $x_0 =$ "good". Individuals are considered "bad" only when they do not take care of their parents whose status label is "good".

Generation t is labeled by a social norm τ and the community informs the succeeding generation $t+1$ of t's status label x_t; that is, the status label of the elderly dependent is known to the breadwinner at period $t+1$. In particular, the breadwinner's action choice is typically a function of the previous generation's status label.

The history of intergenerational transfers may not be known; it becomes known to each generation only to the extent that it is reflected in the status labels of the elderly dependent.[3] A pure (Markov) strategy for a generation is a mapping $s_t : X_{t-1} \to A_t \ \forall t$ specifying a choice of action $s_t(x_{t-1})$ when the previous generation's status label is $x_{t-1} \in X_{t-1}$.

Our equilibrium definition here slightly differs from the last one in the information possessed by the decision makers. Because past actions do not directly affect current or future utility, I do not have to deal with any beliefs regarding the histories.

Definition 4. *A strategy profile $(s_t)_{t=1}^{\infty}$ is a sequentially rational equilibrium if*

$$\forall t, \ \forall x_{t-1} \in X_{t-1}, \ \forall a \in A, \ \ u(s_t, s_{t+1} | x_{t-1}) \geq u(a, s_{t+1} | x_{t-1})$$

[3] Individuals may not know labels of their grandparents or more distant ancestors.

where

$$u(a_t, s_{t+1}|x_{t-1}) = u[a_t, s_{t+1}(\tau(a_t, x_{t-1}))]$$

and

$$u(s_t, s_{t+1}|x_{t-1}) = u[s_t(x_{t-1}), s_{t+1}(\tau(s_t(x_{t-1}), x_{t-1}))].$$

I say that a strategy profile $(s_t)_{t=1}^{\infty}$ is a **tit-for-tat strategy** if

$$s_t(x_{t-1}) = \begin{cases} 1 & \text{if } x_{t-1} = \text{“good”} \\ 0 & \text{if } x_{t-1} = \text{“bad”}. \end{cases}$$

Theorem 3. *In a close-knit community whose social norm is family reciprocity, an efficient level of intergenerational transfers can be sustained as a sequentially rational equilibrium by a tit-for-tat strategy.*

Proof. Each generation is labeled by a social norm τ, i.e., $x_t = \tau(x_{t-1}, a_t)$ for all t. Consider a community whose social norm is family reciprocity, that is, the breadwinner has to take care of his or her parent to be labeled "good" the his parent is "good".

Suppose that each generation plays a tit-for-tat strategy. Take an arbitrary period t. When $x_{t-1} =$ "good", generation t's utility is

$$\begin{aligned} &u(s_t, s_{t+1}|x_{t-1}) = u(1,1) \\ &u(a_t, s_{t+1}|x_{t-1}) = u(a_t, 0) \quad \text{if } a_t \neq 1 \end{aligned}$$

where $u(1,1) > u(a_t, 0)$ for all $a_t \neq 1$. When $x_{t-1} =$ "bad", generation t's utility is

$$\begin{aligned} &u(s_t, s_{t+1}|x_{t-1}) = u(0,1) \\ &u(a_t, s_{t+1}|x_{t-1}) = u(a_t, 1) \quad \text{for all } a_t \end{aligned}$$

where $u(0,1) \geq u(a_t, 1)$. Hence,

$$\forall t,\ \forall x_{t-1} \in X_{t-1},\ \forall a \in A \ \ u(s_t, s_{t+1}|x_{t-1}) \geq u(a, s_{t+1}|x_{t-1}).$$

□

Theorem 3 establishes that in a community whose social norm is family reciprocity, an efficient level of intergenerational transfers can be sustained as a sequentially rational equilibrium by a tit-for-tat strategy. This result is in a sharp contrast with the case without a local community (Theorem 1) where BIGs cannot be supported by pure-strategy equilibrium.

The intuition is as follows. In the basic model, intergenerational cooperation is not sustainable because of the informational advantage of the older generation over the younger generation (Theorem 1). A close-knit community enables intergenerational cooperation because it nullifies the informational advantage of the older generation. Because a status label is a summary statistic of information of unboundedly high orders, it fills the information gap between the older generation and the younger generation so that the older generation cannot manipulate. Thus, by incorporating local communities, this present work demonstrates that the conclusion of the impossibility of intergenerational cooperation under limited information does not persist.

By Theorems 1 and 3 , the intergenerational cooperation is sustained in a close-knit community. Local gossip by neighbors (or extended families) serve to facilitate cooperation between generations.

3.1.4 Erroneous Labeling

In the basic model and community model, I have assumed that the reputation in the community was always correct. That is, it was assumed that neighbors (or extended family) (1) can see how another neighbor treats his or her parent and (2) truthfully report what they observed. However, in reality, one may get a good or bad reputation that he or she does not deserve. In this section, I consider the consequences of relaxing those assumptions.

There are two types of erroneous reputations. One is due to errors in observation. Family affairs may not always be apparent to neighbors, who may circulate gossip based on wrong observations. For example, neighbors may believe that the child is abusing his or her parent even if that is not the case because they happen to observe that the child and the parent are fighting. The other type of erroneous reputation is a result of misreport by neighbors. Even if neighbors view another neighbor's family affairs correctly, they may not truthfully report their observation. For instance, in a community where people tend to speak ill of others, neighbors may gossip that one is abusing his or her parent even if they know that he or she is taking good care of the parent.

Observational Errors

The first type of erroneous labeling is due to observational errors. There may be uncertainty between how much one transfers and how much one appears to transfer. First, even if one takes good care of his or her parent, he or she may appear to neglect the aged parent. For example, an elderly person may die early despite all his or her child's effort. Second, even if one does not take care of his or her parent, he or she may appear to be taking good care of his parent with positive probability. For example, an elderly person may want his or her neighbors to think that he or she has a happy life and thus pretend as if he or she is being taken good care of by the child.

Consider the same community described in Section 3.1.3. Suppose that one may be given a wrong label in such a way that he or she may appear to neglect his or her parent when, in fact, he or she takes care of the parents with probability $1-p$ and he or she may appear to take good care of the parent when he or she does not with probability $1-q$. With this type of erroneous labeling, a variant of family reciprocity in this setting is described as follows.

$$\begin{aligned}\nu_1(x_{t-1}, a_t) &\equiv (\text{Prob}(x_t = \text{“good”}|x_{t-1}, a_t), \text{Prob}(x_t = \text{“bad”}|x_{t-1}, a_t)) \\ &= \begin{cases} (p, 1-p) & \text{if } x_{t-1} = \text{“good” \& } a_t = 1 \\ (1, 0) & \text{if } x_{t-1} = \text{“bad”} \\ (1-q, q) & \text{if } x_{t-1} = \text{“good” \& } a_t \neq 1 \end{cases}\end{aligned}$$

where ν assigns the probabilities that a generation is labeled "good" and "bad" respectively, given the action and the previous generation's label.

Next, I examine whether intergenerational transfers can be sustained in pure Nash equilibrium when observational errors are possible. Suppose that each generation plays a tit-for-tat strategy $(s_t)_{t=1}^{\infty}$. Take an arbitrary period t. When $x_{t-1} =$ "good", generation t's utility is

$$\begin{aligned}u(1, s_{t+1}|x_{t-1}) &= pu(1,1) + (1-p)u(1,0) \\ u(0, s_{t+1}|x_{t-1}) &= (1-q)u(0,1) + qu(0,0).\end{aligned}$$

When $x_{t-1} =$ "bad", generation t's action does change his or her future reputation. Thus, it is optimal for him or her to take $a_t = 0$. Therefore, a tit-for-tat strategy is a Nash equilibrium if and only if

$$\begin{aligned}&u(1, s_{t+1}|x_{t-1} = \text{“good”}) \geq u(0, s_{t+1}|x_{t-1} = \text{“good”}) \\ &\Leftrightarrow p[u(1,1) - u(1,0)] + q[u(0,1) - u(0,0)] \geq u(0,1) - u(1,0). \qquad (1)\end{aligned}$$

In other words, the intergenerational cooperation is sustainable in pure strategy Nash equilibrium when noises $1-p$ and $1-q$ are sufficiently small.

Misreport

The other type of erroneous labeling is due to misreport by neighbors. One may be given a "bad" label even if neighbors know he or she is "good". An example is the society in which people tend to speak ill of others. In addition, one may be given a "good" label even if neighbors know he or she is "bad". A society in which people tend to speak well of others is an example. With this type of erroneous labeling, a variant of family reciprocity in this setting is described as follows.

$$\begin{aligned}\nu_2(x_{t-1}, a_t) &\equiv (\text{Prob}(x_t = \text{“good”}|x_{t-1}, a_t), \text{Prob}(x_t = \text{“bad”}|x_{t-1}, a_t)) \\ &= \begin{cases} (p, 1-p) & \text{if } x_{t-1} = \text{“good” \& } a_t = 1, \text{ or } x_{t-1} = \text{“bad”} \\ (1-q, q) & \text{if } x_{t-1} = \text{“good” \& } a_t \neq 1 \end{cases}\end{aligned}$$

Next, we analyze the sustainability of intergenerational transfers if these types of erroneous reports are possible and social norms are family reciprocity. Suppose that each generation plays a tit-for-tat strategy $(s_t)_{t=1}^{\infty}$. In an arbitrary period t, when x_{t-1} = "good", generation t's utility is the same as that in the case of observational errors. When x_{t-1} = "bad", $a_t = 0$ weakly dominates $a_t \neq 0$. Therefore, a necessary and sufficient condition for a tit-for-tat strategy to be a Nash equilibrium is the same as that in the case of observational errors.[4]

Comparative Statics

The results above have several implications. First, the intergenerational cooperation is sustainable in a pure strategy Nash equilibrium when noises $1-p$ and $1-q$ are sufficiently small. If noise $1-p$ is large, one may think that it is not likely to be recognized if he or she takes good care of his or her parent. Thus, he or she decides not to take care of the parent in equilibrium. With some modification of social norms, more importantly, the maximum amount of sustainable transfers decreases with noise because the noise makes intergenerational cooperation a risky investment. This argument is demonstrated in the following case.

Define payoff functions of the breadwinner and the elderly dependent of the stage game as $v_m : A \to \mathbb{R}$ and $v_o : A \to \mathbb{R}$, respectively. The function v_m is strictly concave, continuously differentiable, and decreasing in a while the function v_o is strictly concave, continuously differentiable, and increasing in

[4] It is not always the case that the case of observational errors and the case of misreport have the same necessary and sufficient condition when social norms are different from a variant of family reciprocity.

a. Let δ denote a discount rate, which is common to all generations. Then, generation t's utility function is denoted as $u(a_t, a_{t+1}) \equiv v_m(a_t) + \delta v_o(a_{t+1})$. Consider another variant of family reciprocity. A social norm τ_{a^*} is **family reciprocity with $\boldsymbol{a^*}$** if

$$x_t = \tau_{a^*}(x_{t-1}, a_t) = \begin{cases} \text{“bad”} & \text{if } x_{t-1} = \text{“good” } \& \ a_t < a^* \\ \text{“good”} & \text{otherwise} \end{cases}$$

and $x_0 =$ “good” where a^* represents the amount of transfers which is considered to be appropriate from a commonsense perspective. For the convenience, consider the case of $a^* > 0$.

For simplicity, suppose that the probabilities of mislabeling are identical, i.e., $p = q$. Then, from (1), intergenerational cooperation is sustainable under family reciprocity with a^* if and only if

$$p \geq \tfrac{u(0,a^*)-u(a^*,0)}{[u(a^*,a^*)-u(a^*,0)]+[u(0,a^*)-u(0,0)]}.$$

Lemma 3. $f(a^*) \equiv \tfrac{u(0,a^*)-u(a^*,0)}{[u(a^*,a^*)-u(a^*,0)]+[u(0,a^*)-u(0,0)]}$ *is continuous and strictly increasing in* a^**.*

Proof. $f'(a^*) > 0$ is equivalent to $\frac{v_m'(a^*)}{v_m(a^*)-v_m(0)} > \frac{v_o'(a^*)}{v_o(a^*)-v_o(0)}$. Since v_m is strictly concave and decreasing, the lefthand side is greater than 1. Similarly, because v_o is strictly concave and increasing, the righthand side is less than 1. Hence, $f'(a^*) > 0$ holds for all $a^* > 0$. □

Lemma 4. *If intergenerational cooperation is sustainable under family reciprocity with* a^**, it is sustainable under any family reciprocity with* $a \leq a^*$*.*

Proof. Take arbitrary p. Because intergenerational cooperation is sustainable under family reciprocity with a^*, $p \geq f(a^*)$. Because $f()$ is strictly increasing (Lemma 3), $f(a^*) > f(a)$ for all $a < a^*$. Hence, $p \geq f(a)$ for all $a < a^*$. □

Proposition 1. *The maximum amount of sustainable transfers* $\hat{a^*}$ $(\neq 0)$ *is increasing in* p

Proof. Take arbitrary $p > f(0)$. Because $f()$ is strictly increasing and continuous (Lemma 3), there exists a unique $\hat{a^*} > 0$ such that $p = f(\hat{a^*})$. Because of Lemma 4, it implies that $\hat{a^*} \equiv \underset{a^*}{\operatorname{argmax}}\{f(a^*) | f(a^*) \leq p\}$. Because f is strictly increasing and continuous (Lemma 3), its inverse function f^{-1} is also strictly increasing and continuous. Hence, the maximum amount of sustainable transfers $\hat{a^*} = f^{-1}(p)$ is increasing in p. □

Thus, Proposition 1 indicates that the maximum amount of sustainable transfers $\hat{a^*}$ ($\neq 0$) is decreasing in noise.

Lastly, given that $1-p$ and $1-q$ are sufficiently small, the probability of erroneous labeling increases with noise $1-p$ on the equilibrium path while not with noise $1-q$. This result indicates that the societies in which people tend to speak well of others support higher welfare than those in which people tend to speak ill of others because a "good" person may be labeled "bad" in the latter societies but not in the former societies in equilibrium.

When the treatment of a "bad" parent is controversial

In some societies, children are encouraged to take care of their parents even if their parents are not good people. Suppose that the treatment of the "bad" parent is controversial and, thus, the label can be "good" or "bad" when one neglects a "bad" parent. This situation is in sharp contrast with the previous setting in which uncertainty in labeling exists regardless of the parent's label.

This community appears to encourage people to be nice to their parents so that they will surely be regarded as "good." Consider a community whose social norm is a variant of family reciprocity. Suppose that one is sometimes given a wrong label in such a way that he or she may be labeled "bad" when he or she does not treat his "bad" parents well. This is an erroneous labeling because, according to family reciprocity, one is supposed to be labeled "good" regardless of one's behavior if the parent is "bad". This community appears to encourage people to be nice to their parents so that they will surely be regarded as good. A variant of family reciprocity in this setting is described as follows.

$$\begin{aligned}\nu_3 &\equiv (\mathrm{Prob}(x_t = \text{“good”}|x_{t-1}, a_t), \mathrm{Prob}(x_t = \text{“bad”}|x_{t-1}, a_t)) \\ &= \tau'(x_{t-1}, a_t) \\ &= \begin{cases} (1,0) & \text{if } (x_{t-1}, a_t) = (\text{“good”}, 1) \text{ or } (\text{“bad”}, 1) \\ (p, 1-p) & \text{if } x_{t-1} = \text{“bad”} \ \& \ a_t \neq 1 \\ (0,1) & \text{if } x_{t-1} = \text{“good”} \ \& \ a_t \neq 1. \end{cases}\end{aligned}$$

According to the social norm ν_3', one is considered "good" with probability 1 when he or she takes care of his or her parents regardless of the parents' label, "good" with probability p when he or she does not take care of the "bad" parents, and "bad" with probability 1 when he or she does not take care of the "good" parents. That is, the community mistakenly regards one as "bad" with probability $1 - p$ when he or she does not take care of the "bad" parents.

Next, I examine whether intergenerational transfers can be sustained in a pure Nash equilibrium in this community. Suppose that each generation plays a tit-for-tat strategy. $(s_t)_{t=1}^{\infty}$. Take an arbitrary period t. When $x_{t-1} =$ "good", no uncertainty exists and it is optimal for generation t to take $a_t = 1$. When $x_{t-1} =$ "bad", generation t's utility is

$$\begin{aligned} u(0, s_{t+1}|x_{t-1}) &= pu(0,1) + (1-p)u(0,0) \\ u(1, s_{t+1}|x_{t-1}) &= u(1,1). \end{aligned}$$

Hence, if and only if $p \geq \frac{u(1,1)-u(0,0)}{u(0,1)-u(0,0)}$, a tit-for-tat strategy is an equilibrium. intergenerational cooperation is sustainable if and only if uncertainty $1 - p$ is sufficiently small. This setting appears to encourage individuals to take good care of their parents because they will surely be labeled "good" as long as they take good care of their parents but they may be mistakenly labeled "bad" if they do not. However, it also discourages individuals from doing so because they expect that their children will be also encouraged to take care of them anyway. This result suggests that the social norms that stress that one should take care of his or her parent may actually discourage individuals from doing so.

3.1.5 Different Norms

The previous section analyzes a community whose social norm is family reciprocity. In this section, I examine two types of social norms in addition to family reciprocity.

Norms of Punishment

According to family reciprocity, an individual does not have to punish his "bad" parents to be labeled "good". When their parents are "bad," an individual will be labeled "good" regardless of his or her behavior. Norms of family reciprocity require only that one should make appropriate amount of transfers when his or her parent is "good". By contrast, some social norms may encourage punishment but not reward. I consider the social norm that requires only that one should NOT make transfers when one's parent is "bad". That is, one does not have to make transfers to the "good" parent to be labeled "good". This social norm is described as follows.

$$\begin{aligned}\chi_t &= (\text{Prob}(x_t = G), \text{Prob}(x_t = B)) \\ &= \tau_2(x_{t-1}, a_t) \\ &= \begin{cases} (1,0) & \text{if } (x_{t-1}, a_t) = (G, a_t) \text{ for all } a_t \text{ or } (B, 0) \\ (0,1) & \text{if } (x_{t-1}, a_t) = (B, a_t) \text{ for all } a_t \neq 0 \end{cases}\end{aligned}$$

With the norm of punishment, the only pure Nash equilibrium is such that no generation makes any intergenerational transfers. Combined with the case of family reciprocity, it suggests that the norm of rewarding is necessary while the norm of punishment is not. As a natural extension, social norms that combine family reciprocity and punishment also support intergenerational transfers in pure strategy Nash equilibrium by a tit-for-tat strategy.[5]

3.2 Three-generations Model

The previous section considered a two-generations model to analyze backward intergenerational transfers. This subsection extends the previous two-generations model to a three-generations model to analyze both backward intergenerational transfers and forward intergenerational transfers, such as school expenses and nutrition provided by individuals to their children. I consider a situation in which a breadwinner decides how much to invest in a backward intergenerational good that benefits only the older dependent and a forward intergenerational good that benefits only the younger dependent. That is, a breadwinner in a family takes care of his or her older dependent and younger dependent. As in the previous subsection, it is demonstrated that neither a BIG nor a FIG is sustained under limited information, and a close-knit community serves to facilitate cooperation between generations.

Consider an infinitely-lived family with three generations alive in each period t ($t = 1, 2, 3 \ldots$). At every period t, a single player, called generation $t + 1$, is born to the family and lives for three periods: t, $t + 1$, and $t + 2$.

[5] The combination of family reciprocity and punishment indicates that one should take care of the "good" parent and punish the "bad" parent:

$$\begin{aligned}\chi_t &= (\text{Prob}(x_t = G | x_{t-1}, a_t), \text{Prob}(x_t = B | x_{t-1}, a_t)) \\ &= \tau(x_{t-1}, a_t) \\ &= \begin{cases} (1,0) & \text{if } (x_{t-1}, a_t) = (G, 1),\ (B, 0) \\ (0,1) & \text{if } (x_{t-1}, a_t) = (G, a_t) \forall a_t \neq 1,\ (B, a_t) \text{ for all } a_t \neq 0\ . \end{cases}\end{aligned}$$

I call generation $t+1$ a younger dependent in period t, a breadwinner in period $t+1$ and an elderly dependent in period $t+2$. In period 1 there are a younger dependent 2 who stays in periods 1, 2, and 3, a breadwinner 1 who stays in periods 1 and 2, and an elderly dependent 0 who stays only for that period.

The breadwinner has positive endowment, and the dependents have endowment that is normalized to zero. Every period t, the breadwinner t decides how much to transfer to his dependents $t+1$ and $t-1$. Let $f_t \in F_t$ denote the amount transferred from the breadwinner t to the younger dependent $t+1$ at period t where action spaces are common across generations, $F_t = F$ for all t. Let $b_t \in B_t$ denote the amount transferred from the breadwinner t to the older dependent $t-1$ at period t where action spaces are common across generations, $B_t = B$ for all t. The commodity is assumed to be infinitely divisible. Both dependents have no choices to make.

Generation t's utility function, $u_t : F_{t-1} \times F_t \times B_t \times B_{t+1} \to \mathbb{R}$, is decreasing in the transfers, f_t and b_t, made in middle-age and increasing in the transfer, f_{t-1}, received at a young age and the transfer, b_{t-1}, received in old age.

That is, a generation's lifetime utility depends only on the actions taken when one is a breadwinner and the actions that the breadwinner takes when he or she is young and old. Utility function $u(\cdot)$ satisfies the condition:

Assumption 3. *$\forall \mathbf{f}, \mathbf{f}' \in F^2, \forall \mathbf{b}, \mathbf{b}' \in B^2$, $u(\mathbf{f}, \mathbf{b}) = u(\mathbf{f}', \mathbf{b}') \;\Rightarrow\; (\mathbf{f}, \mathbf{b}) = (\mathbf{f}', \mathbf{b}')$.*

This condition ensures that $\underset{(f_t, b_t) \in F \times B}{\operatorname{argmax}}\; u(f_{t-1}, f_t, b_t, b_{t+1})$ is unique. Assuming that $\underset{(f_t, b_t) \in F \times B}{\operatorname{argmax}}\; u(f_{t-1}, f_t, b_t, b_{t+1})$ is independent of f_{t-1} and b_{t+1}, I label it $f_t = 0$ and $b_t = 0$, which can be interpreted as no transfers. Note that this overlapping generation game has a unique Markov equilibrium, in which every player chooses 0.

Let $\mathbf{h}_t = ((f_1, b_1), \ldots, (f_{t-1}, b_{t-1}))$ denote the history of preceding actions taken up to period t and $H_t (= F^{t-1} \times B^{t-1})$ denote the set of all possible histories at t. I define the default informational environment as follows:

Definition 5. *For any $t = 1, 2, \ldots$, generation t has* ***m-th order information*** *if he or she knows the actions of the last m generations, $((f_{t-1}, b_{t-1}), \ldots, (f_{t-m}, b_{t-m}))$, but not any action taken prior to $t-m$, $((f_{t-m-1}, b_{t-m-1}), \ldots, (f_1, b_1))$.*

Assumption 4. *There exists a natural number m such that generation t has m-th order information for all $t = 1, 2, \ldots$.*

If (h_t, h'_t) is any pair of histories which differ only in the actions taken by some of players i $(\leq t-m-1)$, then the histories observed by generation t are identical for h_t and h'_t. For example, the first-order information is a record of the preceding generation's past play. That is, the information is limited in the sense that a generation does not know the actions taken by generations prior to the parents in the family.

A pure strategy for generation t is a mapping $s_t : F_{t-m} \times B_{t-m} \times \cdots \times F_{t-1} \times B_{t-1} \to F_t \times B_t$. Let S_t be the set of t's pure strategies. A strategy profile is an infinite sequence $(s_t)_{t=1}^{\infty}$ where $s_t \in S_t$ for all t. Thus, $s_t(f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1})$ is the element of $F_t \times B_t$ which is induced by the observed history $(f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1})$ is played.

An action by generation t, (f_t, b_t), when he or she observes $(f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1})$, induces an observed history for $t+1$, $(f_{t-m+1}, b_{t-m+1}, \ldots, f_t, b_t)$. Different actions by generation t generate different information for $t+1$, and generation $t+1$ varies his or her actions with the observed history, $(f_{t-m+1}, b_{t-m+1}, \ldots, f_t, b_t)$.

As in the two-generations model, I define our equilibrium definition as follows.

Definition 6. *A strategy profile $(s_t)_{t=1}^{\infty}$ is a sequentially rational equilibrium if $\forall t$, $\forall (f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1}) \in F_{t-m} \times B_{t-m} \times \cdots \times F_{t-1} \times B_{t-1}$, $\forall f_t \in F_t$, $\forall b_t \in B_t$, $u(s_t, s_{t+1} | f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1}) \geq u(f_t, b_t, s_{t+1} | f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1})$ where*

$$u(a_t, s_{t+1} | f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1}) = u[a_t, s_{t+1}(f_{t-m+1}, b_{t-m+1}, \ldots, f_t, b_t)]$$

and

$$u(s_t, s_{t+1} | f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1}) = u[s_t(f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1}), s_{t+1}(f_{t-m+1}, b_{t-m+1}, \ldots, f_t, b_t)]$$

The following theorem is a variant of Theorem 1 in **?** and Theorem 1.

Theorem 4. *Under Assumption 4, the overlapping generations game with m-th order information has a unique pure strategy equilibrium where no intergenerational transfer is made.*

Proof. Suppose that generation t's equilibrium strategy s_t is conditioned on action taken by generation $t-m$. Then, there exist $f_{t-m}, b_{t-m}, f'_{t-m}, b'_{t-m}$ such that $u_t(s_t, s_{t+1} | f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1}) > u_t(s_t, s_{t+1} | f'_{t-m}, b'_{t-m}, \ldots, f_{t-1}, b_{t-1})$. Because s_{t+1} is conditioned only on actions taken by generations $t-m+1, \ldots, t$ and is independent of generation $t-m$'s action, generation t can improve his or her payoff by choosing $s_t(f_{t-m}, b_{t-m}, \ldots, f_{t-1}, b_{t-1})$ instead of $s_t(a'_{t-m}, \ldots, a_{t-1})$ when he or she observes $s_t(f'_{t-m}, b'_{t-m}, \ldots, f_{t-1}, b_{t-1})$. Hence, generation t's equilibrium strategy does not condition on $t-m$'s actions. Similarly, generation t's equilibrium strategy is not conditioned on k's action for any $k < t$. Thus, generation t's equilibrium strategy is not conditioned on history. Therefore generation t's best response is $f_t = b_t = 0$, which forms a unique pure strategy equilibrium where no intergenerational transfer is made. $\square$

Theorem 4 indicates that when individuals have limited information about past events, no strategy profile (that does not have to be Markov) in which players condition their behavior on the observed history, which is payoff-irrelevant, constitutes an equilibrium. Thus, again, intergenerational cooperation cannot be supported in pure-strategy equilibrium when information about past events is limited.

3.2.1 Local Communities

As in the two-generations model, I modify the basic three-generations model by adding a local community that provides information about the history of play in each family through gossiping. I consider a situation in which families in the previous basic three-generations model reside in a close-knit community. In a close-knit community, the breadwinner receives a reputation based on how he or she treated his or her parent and child, and that reputation will become known to the child through neighbors' gossip. The transfers he or she will receive from the child in old age may depend on that reputation, which in turn determines the child's reputation. In this setting, transfers can be supported by pure-strategy equilibria. Formally, this situation is modeled as follows.

Before introducing a local community to the basic model, I define the "efficient" level of transfers. Let $\underset{f\in F\, b\in B}{\operatorname{argmax}}\, u((f,t),(f,t)) = (1,1)$ be the efficient transfers, which I assume to be different from $(0,0)$, the individually optimal action. Otherwise, it is trivial that no transfer is sustained in equilibrium. From Assumption 3,

Lemma 5. $u((0,0),(1,1)) > u((1,1),(1,1)) > u((0,0),(0,0)) > u((1,1),(0,0))$

In each period t, the breadwinner t is assigned a status label x_t of a finite set $X_t = \{\text{“good”}, \text{“bad”}\}$. A generation's status label $x_t \in X_t$ is determined through $\tau : X_{t-1} \times F_t \times B_t \to X_t$. τ specifies the status label of generation t in the next period, $\tau(x_{t-1}, f_t, b_t) \in X_t$, when the previous generation's status label is $x_{t-1} \in X_{t-1}$ and one's current actions are $f_t \in F_t$ and $b_t \in B_t$. A social norm, τ, is common knowledge.

A social norm τ is **family reciprocity** if

$$x_t = \tau(x_{t-1}, f_t, b_t) = \begin{cases} \text{“bad”} & \text{if } x_{t-1} = \text{“good” } \& \ (f_t \neq 1 \text{ or } b_t \neq 1) \\ \text{“good”} & \text{otherwise} \end{cases}$$

and x_0 = “good”. Individuals are considered “bad” only when they did not take care of their parents or children when their parents' status label is “good”.

Generation t is labeled by a social norm τ and the community informs the succeeding generation $t+1$ of t's status label x_t; that is, the status label of the elderly dependent is known to the breadwinner at period $t+1$. In particular, the breadwinner's action choice is typically a function of the previous generation's status label.

The history of intergenerational transfers may not be known; it becomes known to each generation only to the extent that it is reflected in the status labels of the elderly dependent. A pure (Markov) strategy for a generation is a function $s_t : X_{t-1} \to F_t \times B_t \ \forall t$ specifying a choice of action $s_t(x_{t-1})$ when the previous generation's status label is $x_{t-1} \in X_{t-1}$.

Our equilibrium definition here slightly differs from the last one in the information possessed by the decision makers. Because past actions do not directly affect current or future utility, I do not have to deal with any beliefs regarding the histories.

Definition 7. *A strategy profile* $(s_t)_{t=1}^{\infty}$ *is a sequentially rational equilibrium if*

$$\forall t,\ \forall x_{t-1} \in X_{t-1},\ \forall f \in F,\ \forall b \in B,\ \ u(s_t, s_{t+1}|x_{t-1}) \geq u(f, b, s_{t+1}|x_{t-1})$$

where

$$u(f_t, b_t, s_{t+1}|x_{t-1}) = u[f_t, b_t, s_{t+1}(\tau(f_t, b_t, x_{t-1}))]$$

and

$$u(s_t, s_{t+1}|x_{t-1}) = u[s_t(x_{t-1}), s_{t+1}(\tau(s_t(x_{t-1}), x_{t-1}))].$$

I say that a strategy profile $(s_t)_{t=1}^{\infty}$ is a **tit-for-tat strategy** if

$$s_t(x_{t-1}) = \begin{cases} (1,1) & \text{if } x_{t-1} = \text{“good”} \\ (0,0) & \text{if } x_{t-1} = \text{“bad”}. \end{cases}$$

Theorem 5. *In a close-knit community whose social norm is family reciprocity, an efficient level of intergenerational transfers can be sustained as a sequentially rational equilibrium by a tit-for-tat strategy.*

Proof. Each generation is labeled by a social norm τ, i.e., $x_t = \tau(x_{t-1}, f_t, b_t)$ for all t. Consider a community whose social norm is family reciprocity; that is, the breadwinner has to take care of both his or her parent and child to be labeled “good” if his parent is “good”.

Suppose that each generation plays a tit-for-tat strategy. Take an arbitrary period t. When $x_{t-1} =$ “good”, generation t's utility is

$$\begin{aligned} &u(s_t, s_{t+1}|x_{t-1}) = u(1,1) \\ &u(f_t, b_t, s_{t+1}|x_{t-1}) = u(f_t, b_t, 0) \quad \text{if } f_t \neq 1 \text{ or } b_t \neq 1 \end{aligned}$$

where $u((1,1),(1,1)) > u((f_t, b_t),(0,0))$ for all $(f_t, b_t) \neq (1,1)$. When $x_{t-1} =$ “bad”, generation t's utility is

$$\begin{aligned} &u(s_t, s_{t+1}|x_{t-1}) = u(0,1) \\ &u(f_t, b_t, s_{t+1}|x_{t-1}) = u(f_t, b_t, 1) \quad \text{for all } (f_t, b_t) \neq (1,1) \end{aligned}$$

where $u((0,0),(1,1)) \geq u((f_t, b_t),(1,1))$. Hence,

$$\forall t,\ \forall x_{t-1} \in X_{t-1},\ \forall (f_t, b_t) \neq (1,1),\ \ u(s_t, s_{t+1}|x_{t-1}) \geq u(f, b, s_{t+1}|x_{t-1}).$$

□

Theorem 5 establishes that in a community whose social norm is family reciprocity, an efficient level of intergenerational transfers can be sustained as a sequentially rational equilibrium by a tit-for-tat strategy. This result contrast sharply with the case without a local community (Theorem 4) in which any transfers cannot be supported by pure-strategy equilibrium. Note that, in two-generations model, BIGs alone without FIGs can be supported in a community whose social norm is the previous family reciprocity.

By Theorems 4 and 5 , the intergenerational cooperation is sustained in a close-knit community. Depending on social norms, there are three types of equilibria: one type is that no intergenerational transfers are made; another is that only BIGs are made; the other is that both BIGs and FIGs are made. For FIGs to be sustainable, BIGs are needed because they give a breadwinner incentives to make transfers. Local gossip by neighbors (or extended families) serve to facilitate cooperation between generations.

4 Empirical Results

In Section 3, I present formal models that show that intergenerational family transfers such as nursing care (from children to parents) or education expenditures (from parents to children) increase as local networks become stronger. Empirical results in this section support this argument: Individuals who interact with their neighbors tend to provide good nursing care to their aged parents and spend more on their children's education. Furthermore, assuming that the welfare provision by the government and support from the family are substitutable at individual-level preference, my theory suggests that individuals who interact with their neighbors may demand less from the government, an implication that is supported by the empirical findings. The last subsection suggests that the regional difference in household structures between northeast and southwest Japan is due to regional diversity in neighborhood interaction.

4.1 Nursing Care and Interaction with Neighbors

I use the cross-sectional survey data *Kongono Seikatsuni Kansuru Anketo* for 1999 to measure how the level of nursing cares that individuals intend to provide to their aged parents varies and how such attitudes correlate with the level of their interactions with their neighbors in Japan.

My dependent variable here is a scale (1-8) for how the respondents intend to care for their own parents. Individuals respondeded on a 1-8 scale, with 1 being "I will take care of my aged parent at home by myself", 2 being "I will take care of my aged parent at home with support from other relatives", 3 being "I will take care of my aged parent at home with outside support", 4 being "Mainly other relatives will take care of my aged parent and I will visit them for help", and so on.

The key explanatory variable is the level of the respondents' interactions with their neighbors. Individuals responded on a $1-3$ scale, with 1 indicating "friendly communicate with neighbors", 2 indicating "only say "hi" to neighbors", and 3 indicating "do not interact with neighbors much".

Table 1 shows that the level of nursing care intended for the aged parents is positively associated with the level of interaction with neighbors. Those individuals who frequently interact interact with their neighbors tend to provide a higher level of nursing care for their parents.

Table 1: Determinants of nursing care at the individual level

Interaction with neighbors at individual level	**0.256***	**(0.074)**
Have a child	0.492*	(0.112)
Male	-0.316*	(0.084)
Age	-0.006	(0.003)
Income	0.015	(0.02)
N	2098	

* $p \leq 0.05$. Entries are logistic regression estimates with standard errors in parentheses.

Remittance and Participation in Home Country Organizations

Table 2 shows that the decision about whether the immigrants send money to their home countries is positively associated with whether they participate in any kind of organization that is associated with their home countries, even after controlling for their attachment to their home countries.

Table 2: Determinants of remittance from immigrants to their home countries

Participation in an organization which is associated with home country	**0.830*** **	**(0.136)**
Home country feels like home to R rather than US	**0.833*** **	**(0.189)**
Income	0.0599***	(0.0180)
Number of Children	0.0253**	(0.0093)
Age	-0.0199*	(0.0081)
Male	0.0024	(0.0839)
Asian	-0.537***	(0.0876)
White, non-Hispanic	-1.10***	(0.192)
Black, non-Hispanic	0.169	(0.373)
N	2721	

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1

Entries are logistic regression estimates with standard errors in parentheses.

4.2 Education Expenditures and Local Activities

As well as the backward intergenerational transfers such as nursing care, my models suggest that the forward intergenerational transfers such as education are also sustainable with local gossiping.

I use Japanese General Social Survey data from 2006 to measure how the level of household expenditure on children's education varies and how such behavior correlates with the level of parents' participation in local activities in Japan.

My dependent variable here is a scale (1-6) for how much the respondent spend on their children compared to the average level. Individuals respond on a 1-6 scale with 1 being "far less than the average", 2 being "less than the average", 3 being "about the average", an so on.

The key explanatory variable is a number of local activities that the respondents participated in. The respondents were asked three questions: whether they participated in a community cleaning, whether they participated in community policing, and whether they participated in recycling activities in their community. The variable is assigned 3 when the respondent participated in all the activities above and 0 when the respondent participated in none of them.

Table 3 shows that household education expenditures are positively associated with the level of respondents' participation in three local activities: the community patrol, cleaning of the community, and a recycling campaign in the community. The respondents who participated in more activities tend to spend more on their children's educations.

Table 3: Determinants of household education expenditures

The number of local activities that R participates in	**0.147***	**(0.056)**
Community size	-0.07	(0.054)
Male	0.231*	(0.102)
Age	0.03*	(0.004)
Married	0.249	(0.165)
Number of children	-0.332*	(0.070)
Household income level	0.468*	(0.057)
N	1483	

* $p \leq 0.05$. Entries are logistic regression estimates with standard errors in parentheses.

4.3 Political Opinion

By implication, my theory suggests that individuals in a close-knit community demand less from the government. According to my theory, individuals in a close-knit community are or are expected to be well-taken care of by their children. Assuming that welfare provision by the government and support from the family are substitutable at the individual preference level, those individuals in a close-knit community will need less from the government compared to those who cannot expect much support from their children.

Empirical results support this implication. I use Japanese General Social Survey data from 2006 to measure how individual attitudes towards social spending vary and how such attitudes correlate with the level of their participation in local activities in Japan.

First, I examine the implication about the backward intergenerational goods. My dependent variable is a scale (1-3) for how much the respondent supports the government spending on social security. The question asked was "What do you think about the government expenditure on social security?" and individuals responded on a scale of $1-3$, with 1 being *It is too much* and 3 being *It is too little*.

Table 4 shows a significant association between the level of respondents' participation in local activities and their opinion about the government spending on social security. The respondents who actively participated in local activities such as the community patrol, cleaning of the community, and a recycling campaign in the community tended to think that the government was spending too much on social security.

Second, I examine the implication about the forward intergenerational goods. My dependent variable is a scale (1-3) for how much the respondents supported the government expenditure on education. The question asked was "What do you think about the government expenditure on education?" and individuals responded on a scale of $1-3$, with 1 being *I tis too much* and 3 being Ît is too little.

Table 5 shows a significant association between the level of respondents' participation in local activities and their opinions about the government spending on education. The respondents who actively participated in local activities such as the community patrol, cleaning of the community, and a recycling campaign in the community tended to think that the government was spending too much on social security.

Therefore, Table 4 and Table 5 indicate that individuals who actively

Table 4: Determinants of R's opinion about the government spending on social security

The number of local activities that R participates in	**-0.137***	**(0.054)**
Community size	-0.073	(0.049)
Male	-0.702*	(0.095)
Age	0.008*	(0.003)
Married	-0.451*	(0.12)
Number of children	0.148*	(0.053)
Household income level	-0.154*	(0.051)
N	1984	

* $p \leq 0.05$. Entries are logistic regression estimates with standard errors in parentheses.

Table 5: Determinants of R's opinion about the government spending on education

The number of local activities that R participates in	**-0.108***	**(0.049)**
Community size	0.018	(0.044)
Male	-0.704*	(0.085)
Age	0.011*	(0.003)
Married	-0.278*	(0.111)
Number of children	0.032	(0.049)
Household income level	-0.081	(0.046)
N	1984	

* $p \leq 0.05$. Entries are logistic regression estimates with standard errors in parentheses.

participated in local activities tend to demand less from the government.

4.4 Northeast Japan vs. Southwest Japan

Finally, the empirical findings also suggest that the apparently regional difference in family-based welfare provision is due to their variation in the community networks. As sociologists argue, individuals tend to live in households consisting of more than two generations in northeast Japan compared to their counterparts in southwest Japan. The following empirical results suggest the seeming difference in household structures between northeast and southwest Japan are due to their variation in their community networks.

In this subsection, I use the cross-sectional survey data *Kongono Seikatsuni Kansuru Anketo* from 1997 to measure how individuals' living arrangement varies across these regions in Japan and how such arrangements are associated with the level of their interactions with their neighbors.

The dependent variable in my analysis is a 0-1 dummy which responds to the question, "Does your household consist of three or more generations?". There are two explanatory variables of interest. One is a 0-1 dummy for whether the respondent lives in northeast Japan or not. Because the sample consists of individuals who reside in northeast or southwest Japan, resonses indicating southwest Japan were assigned the dummy variable 0. The other explanatory variable of interest is the level of the respondents' interaction with their neighbors. Individuals respond on a $1-3$ scale with 1 being "friendly communication with neighbors", 2 being "only say "hi" to neighbors", and 3 being "do not interact with neighbors much".

The second set of results in Table 6 are for a model in which the individual-level variable of neighborhood interaction is added to the first model as a predictor. In the first model in Table 6, individuals who reside in northeast Japan tend to be members of households consisting of three or more generations. However, the second model shows that the positive association between northeast residence and three-generations households becomes insignificant, controlling for individual-level neighborhood interaction.

In addition to the logistic regression of household structures, Table 7 indicates an ordered logistic regression morel whose response variable is individual-level neighborhood interaction. It shows that individuals in northeast Japan tend to interact with their neighbors more frequently than did their counterparts in southwest Japan.

Those results suggest that the apparent regional difference in household structures is due to the regional variation in neighborhood interactions. Put it differently, individuals in northeast Japan appear to tend to be members of households consisting of more than two generations because they tend to more frequently interact with their neighbors more frequently.

Table 6: Determinants of the three or more generations living together

	Model 1	Model 2
Northeast (dummy)	**0.38***	**0.332**
	(0.176)	**(0.178)**
Female	0.237	0.203
	(0.176)	(0.177)
Age	-0.016*	-0.02*
	(0.006)	(0.007)
Household Income	0.071*	0.075*
	(0.027)	(0.028)
House ownership	2.59*	2.523*
	(0.433)	(0.435)
Level of neighborhood	—	**0.168***
interaction		**(0.06)**
N	712	712

* $p \leq 0.05$. Entries are logistic regression estimates with standard errors in parentheses. Both models include individual-specific intercepts (not shown).

Table 7: Determinants of the individual level of neighborhood interaction

Northeast (dummy)	**0.395***	(0.149)
Female	0.426*	(0.148)
Age	0.029*	(0.006)
Household Income	-0.01	(0.023)
House ownership	0.877*	(0.214)
N	700	

* $p \leq 0.05$. Entries are logistic regression estimates with standard errors in parentheses.

	Estimate	Std. Error	z value	Pr(>|z|)
Northeast Dummy	-0.24	0.18		
Female	0.18	0.10	1.78	0.07
F2_1	-0.01	0.00	-2.60	0.01
as.numeric(F7)	0.07	0.02	4.24	0.00
as.numeric(V139)	2.48	0.22	11.15	0.00
municipality	-0.00	0.00	-3.37	0.00

5 Conclusion

This paper investigated how the family works as a provider of social protection. More specifically, it argues that differences in the level of transfers within the family are due to demographic variation in community networks. To analyze the sustainability of intergenerational transfers, I propose game theoretical models of overlapping generations in which the breadwinner makes transfers to the parent and child. A novel feature of the models is that there is a local community that may supply information about its members' past behavior. I demonstrate that intergenerational transfers can be sustained if neighbors gossip about each other. Furthermore, I also demonstrate that the maximum amount of sustainable transfers decreases as labeling becomes noisier. By implication, my theory suggests that individuals in a close-knit community prefer lower levels of social protection. Empirical results from Japan support this argument: Individuals who frequently interact with their neighbors tend to provide better nursing care to their aged parents, spend more on their children's education, and demand less from the government than those who do not interact frequently with their neighbors. Furthermore, it is suggested that the apparent regional difference in household structures is due to the regional variation in neighborhood interactions. That is, individuals in northeast Japan appear to tend to be members of households consisting of more than two generations because they tend to interact with their neighbors more frequently.