4-14-2013

Stochastic dominance and demographic policy evaluation: a critique

Juan Carlos Cordoba
Iowa State University, cordoba@iastate.edu

Xiying Liu
Iowa State University, xiyingl@iastate.edu

Stochastic dominance and demographic policy evaluation: a critique

Abstract
Stochastic dominance (SD) is commonly used to rank income distribution and assess social policies. The literature argues that SD is a robust criterion for policy evaluation because it requires minimal knowledge of the social welfare function. We argue that, on the contrary, SD is not a robust criterion. We do this by carefully introducing microfoundations into a model by Chu and Koo (1990) who use SD to provide support to family-planning programs aiming at reducing the fertility of the poor. We show that fertility restrictions are generally detrimental for both individual and social welfare in spite of the fact that SD holds. Our findings are an application of the Lucas' Critique.

Keywords
fertility, welfare, income distribution, children, demographic policies, one child policy, stochastic dominance

Disciplines
Economics

Stochastic Dominance and Demographic Policy Evaluation: A Critique

Juan Carlos Cordoba, Xiying Liu

Working Paper No. 13009
April 2013

IOWA STATE UNIVERSITY
Department of Economics
Ames, Iowa, 50011-1070

Stochastic Dominance and Demographic Policy Evaluation: A Critique

Juan Carlos Cordoba* and Xiying Liu†

April 2013

Abstract

Stochastic dominance (SD) is commonly used to rank income distribution and assess social policies. The literature argues that SD is a robust criterion for policy evaluation because it requires minimal knowledge of the social welfare function. We argue that, on the contrary, SD is not a robust criterion. We do this by carefully introducing microfoundations into a model by Chu and Koo (1990) who use SD to provide support to family-planning programs aiming at reducing the fertility of the poor. We show that fertility restrictions are generally detrimental for both individual and social welfare in spite of the fact that SD holds. Our findings are an application of the Lucas' Critique.

Keywords: fertility, children, welfare, demographic policies, one child policy, income distribution, stochastic dominance.

JEL classification: I10, I31, J17, O57

*Iowa State University

†Iowa State University

1 Introduction

A classical literature on the measurement of inequality claims that stochastic dominance provides a robust criterion to rank income distributions. This literature originated in papers by Kolm (1969) and Atkinson (1970), and was extended by Dasgupta et al. (1973), Rothschild and Stiglitz (1973), Saposnik (1981, 1983), and Foster and Shorrocks (1988a, 1988b) among many.[1] As summarized by Foster and Shorrocks (1988a), first order stochastic dominance (FSD) “can be regarded as the welfare ordering that corresponds to unanimous agreement among all monotonic utilitarian functions." As such, FSD seemingly provides a robust criterion for policy evaluation because it only requires minimal knowledge of the social welfare function. A natural prescription of this literature would be to look for policies that improve the distribution of incomes in the FSD sense.

An important application of stochastic dominance is the one by Chu and Koo (1990) (CK henceforth). They use FSD to evaluate the consequences of changing the reproduction rate of a particular income group. Using a Markovian branching framework with differential fertility among income groups, they show that an exogenous reduction in the fertility of the poor results in a sequence of income distributions that *conditionally* first-degree stochastically dominate (CFSD) the original distribution. CFSD implies FSD. CK argue that stochastic dominance “provides us with very strong theoretical support in favor of family-planning programs that encourage the poor in developing countries to reduce their reproductive rate (pp. 1136)." Numerical simulations of CK's model further confirm that more general fertility reduction programs that disproportionately targets lower income groups, such as the *One Child Policy*, or policies that promote fertility of high income groups, should increase social welfare.[2] These policies generally result in a sequence of income distributions that dominates

[1] A more complete list of references can be found in Davidson and Duclos (2000) and Atkinson and Brandolini (2010). A more precise terminology is "welfare dominance" as used by Foster and Shorrocks (1988b). We use stochastic dominance because this is the term used in the paper that is the focus of our critique.

[2] On policies seeking to increase the fertility of high income groups, the New York Times reports about the Chinese policy of “upgrading" the quality of their population in order to increase its international competitiveness. It suggests an strategy that includes stigmatizing unmarried women older than 28, who are typically highly educated, as “leftover" women. See http://www.nytimes.com/2012/10/12/opinion/global/chinas-leftover-women.html? Last accessed 3/15/2013

the distribution without the program in the first order stochastic sense.

CK's results are nevertheless puzzling. Basic economic principles suggest that absent externalities or market failures individuals' decisions should be efficient. In fact, various authors have shown that fertility choices made by altruistic parents, i.e. parents who care about the number and welfare of their children, are socially optimal under certain conditions. Early papers in this category include Pazner and Razin (1980), Willis (1985), Becker (1983), Eckstein and Wolpin (1985). Recent work by Golosov et al. (2007) further shows that market allocations are Pareto optimal in a variety of models of endogenous fertility. These findings suggest that family planning programs aiming at reducing the fertility of the poor do not necessarily have the strong theoretical support claimed by CK. Lam (1993, pp 1043) expresses similar skepticism.

Unfortunately CK do not fully spell out the decision problem of individuals, a common feature of the literature cited in the first paragraph. However their two main assumptions, grounded on empirical evidence, are in fact hard to rationalize by frictionless models of fertility. First, they assume intergenerational mobility across income and consumption groups but complete market models, such as the Barro-Becker model, predict no mobility.[3] Second, they assume that fertility decreases with individual income, a feature that is also difficult to rationalize by efficient models of fertility (see Cordoba and Ripoll, 2010). It is possible that behind these two assumptions there are some implicit frictions explaining why fertility is suboptimal in CK's model and intervention is welfare enhancing.

This paper revisits the question of optimality of family planning programs as envisioned by CK but explicitly taking into consideration the household decision problem. For this purpose we use a version of the Barro and Becker (1988, 1989) model enriched to study issues of income distribution. Individuals in our model differ in their innate abilities, are altruistic toward their descendants, and choose their own fertility optimally. Abilities are random, determined at birth and correlated with parental abilities. Insurance markets are available

[3] Mobility is still hard to obtain by models of incomplete markets. For example, Alvarez (1999) finds lack of mobility in the Barro-Becker model even in the face of uninsurable idiosyncratic risk. Using a non-altruistic framework, Raut (1990) also finds that the economy reaches a steady state, with no mobility, in two periods.

but parents cannot leave negative bequests to their children. Due to the assumed market incompleteness, mobility arises in equilibrium and fertility differs across ability groups.

The equilibrium of the model satisfies the two assumptions postulated by CK. First, fertility decreases with ability in the presence of uncertainty about children's abilities. To the extent of our knowledge, this result is novel and of independent interest by itself. Although there is a literature documenting and studying a negative relationship between fertility and ability, obtaining such negative relationship within a fully dynamic altruistic model with uncertainty is novel.[4] The negative relationship arises from the interplay of two opposites forces. On the one hand, higher ability individuals face a larger opportunity cost of having children due to the time cost of raising children. On the other hand, higher ability individuals enjoy a larger benefit of having children when abilities are intergenerationally persistent. We find that the effect of ability on the marginal cost dominates its effect on the marginal benefit if the intergenerational persistence of abilities is not perfect. This explains why fertility decreases with ability. Second, the equilibrium of the model exhibits mobility. In particular, the equilibrium is characterized by a Markov branching process satisfying the *Conditional Stochastic Monotonicity* property. This requirement means that if a kid from a poor family and a kid from a rich family both fall into one of the poorest classes, it is more likely that the poor kid will be poorer than the rich kid.

Given that the equilibrium of our model satisfies the assumptions postulated by CK, direct application of their Theorem 2 implies that a reduction in the fertility of the poor generates a sequence of income distributions that dominates the original distribution in all periods in the first order stochastic sense. In particular, average income and consumption increase for all periods. This result comes from two forces. First, average ability of (born) individuals increases because the poor have proportionally more low ability children as a result of the assumed conditional stochastic monotonicity property. Second, consumption and income of the poor strictly increases because they spend less time and resources raising

[4] See for example Becker (1960), Jones and Tertilt (2006), Kremer (1993), Hansen and Prescott (2002), Cordoba and Ripoll (2010).

children. However, contrary to CK's claim, we find that individual and social welfare fall. Our main result, Proposition 8 shows that fertility restrictions of any type, not only for the poor, unequivocally reduce individual and social welfare in our model, in spite of the strong degree of market incompleteness. Hence we conclude that stochastic dominance alone is not a sound criterion to rank social welfare as claimed by Chu and Koo in particular, and by a larger literature in general.

The primary reason why stochastic dominance fails to rank welfare properly is because it does not take into account the fact that indirect utility functions are not invariant to the policies in place. As we show, a policy that restricts fertility in our model reduces the set of feasible choices and invariably reduces welfare of all individuals in all generations, even those whose fertility is not directly affected. This is because altruistic parents care not only about their own consumption and fertility but also care about the consumption and fertility of all their descendants. Furthermore, the welfare of those individuals who are not born under the new policy also falls, or at least does not increase. Social welfare falls because the welfare of all individuals, born and unborn, either falls or remain the same. This is the case, for example, if social welfare is defined as classical (Bentham) utilitarianism, a weighted sum of the welfare of all present and future individuals. The result also holds for versions of classical utilitarianism that are consistent with the Barro-Becker concept of diminishing altruism. An interpretation of our results is that the positive effect on welfare of fertility restrictions, namely higher average consumption, is dominated by the negative effect of an smaller dynasty size.

CK define social welfare as average (Mills) utilitarianism rather than classical utilitarianism. Under this definition, social welfare can increase even if the welfare of all individuals falls if population falls even more. The net effect of fertility restrictions on social welfare depends in this case on the relative strength of two opposite forces. On the one hand, the *distribution* of abilities and incomes improves for all periods, as stressed by CK. On the other hand, the welfare of all individuals fall. Propositions 9 and 10 provide two examples in which the later force dominates and social welfare, defined as average welfare, falls not only in present value but also for all periods. These are counterexamples to the claim that stochastic dominance is a sufficient condition to rank social welfare, even when welfare is defined as average utilitarianism. We further provide a variety of numerical simulations to illustrate that our results are general, not just extreme examples.

Our results challenge the policy implications of CK's paper but also the broader literature, mentioned in the first paragraph, claiming that stochastic dominance alone provide robust normative implications. We show that carefully modeling the microfoundations of the problem makes a difference and can reverse the conclusions obtained by simple stochastic dominance criteria. Our findings are an application of the Lucas' critique. CK's results are based on the assumption that reduced form parameters and indirect utility functions are invariant to policy changes. Specifically, the fertility rate as well as the indirect utility functions of individuals are assumed to be invariant the policies in place. However these are not structural parameters but function of deeper parameters, those governing preferences, technologies and policies in place. Policy evaluations based on the assumed constancy of the parameters may be misleading. In his classic critique, Lucas argued that the observed negative relationship between unemployment and inflation cannot actually be exploited by policymakers to systematically reduce unemployment. The analogous argument in our context is that the observed negative relationship between fertility and income cannot be exploited by policymakers to improve social welfare.

In addition to the papers already mentioned, our paper is related to Alvarez (1999). He studies an economy with idiosyncratic shocks, incomplete markets and endogenous fertility choices by altruistic parents. Our endowment economy is a version of his model, one with non-negative bequest constraints. In equilibrium no individual leaves positive bequests. This is a stronger degree of market incompleteness than that in Alvarez and it explains why mobility arises in the equilibrium of our model but not in his. As a result, our model maps exactly into CK's Markovian model. There is a related literature that studies fertility policies in general equilibrium model. A recent example is Liao (2013) who studies the *One Child Policy* in a model with human capital accumulation and she arrives similar conclusions. Our paper is complementary to hers. Our framework is simpler in that we study an endowment economy but richer in the heterogeneity and its ability to generate mobility. This allow us to study in detail the dynamics of the income distribution and the soundness of stochastic dominance concepts to rank social welfare.

The rest of the paper is organized as follows. Section 2 revisits the basic connection between fertility, distribution of income and social welfare in models with exogenous fertility. The section reviews the result of CK and provides further analysis. Section 3 endogenizes fertility and shows that fertility generally decreases with ability and income. Section 4 studies social policies. It shows the basic limitation of CK's assumptions and argues that fertility policies typically reduces social welfare. Numerical simulations and robustness checks are performed in this section. Section 5 concludes. Proofs are in the Appendix.

2 Distribution and Social Welfare with Exogenous Fertility

Consider an economy populated by a large number of individuals who live for one period. Individuals differ in their labor endowments, or earning abilities. Let $\Omega \equiv \{\omega_1, \omega_2, ..., \omega_n\}$ be the set of possible abilities, where $0 < \omega_1 < ... < \omega_n$. The technology of production is linear in ability: one unit of labor produces one unit of perishable output. In this section, the income of an individual is equal to his/her ability. Let $f(\omega)$ be the fertility rate of an individual with ability ω. It satisfies the following assumption.

Assumption 1. $f(\omega_i)$ is decreasing in ability.

2.1 Abilities

Ability is determined at birth and correlated with the ability of the parent. Ability is drawn from the Markov chain M where $M_{ij} = \Pr(\omega_{\text{child}} = \omega_i | \omega_{\text{parent}} = \omega_j)$ for ω_i and $\omega_j \in \Omega$. As in CK, assume that M satisfies the following condition:

Assumption 2. Conditional Stochastic Monotonicity (CSM):

$$\frac{\sum_{i=1}^{I} M_{i1}}{\sum_{j=1}^{J} M_{j1}} \geq \frac{\sum_{i=1}^{I} M_{i2}}{\sum_{j=1}^{J} M_{j2}} \geq ... \geq \frac{\sum_{i=1}^{I} M_{in}}{\sum_{j=1}^{J} M_{jn}},\ 1 \leq I \leq J \leq n$$

Assumption 2 means that if a poor kid and a rich kid both fall into one of the poorest classes, it is more likely that the poor kid will be poorer than the rich kid. Assumption 2 assures intergenerational persistence of abilities: higher ability parents are more likely to have higher ability children. CSM implies first order stochastic dominance. To see this notice that when $J = n$ the condition becomes:

$$\sum_{i=1}^{I} M_{i1} \geq \sum_{i=1}^{I} M_{i2} \geq ... \geq \sum_{i=1}^{I} M_{in},\ 1 \leq I \leq n.$$

Two examples of Markov chains satisfying Assumption 2 are an i.i.d. process and quasi-diagonal matrices of the form:

$$M_1' = \begin{bmatrix} a+b & c & 0 & 0 & .. & 0 & 0 \\ a & b & c & 0 & .. & 0 & 0 \\ 0 & a & b & c & .. & 0 & 0 \\ .. & .. & .. & .. & .. & 0 & 0 \\ 0 & 0 & 0 & 0 & a & b & c \\ 0 & 0 & 0 & 0 & 0 & a & b+c \end{bmatrix}.$$

where $(a, b, c) \gg 0$, $a + b + c = 1$ and $b > 0.5$.

We further assume that M has a unique invariant distribution, μ, where μ satisfies:

$$\mu(\omega_j) = \sum_{\omega_i \in \Omega} \mu(\omega_i) M(\omega_j, \omega_i) \text{ for all } \omega_j \in \Omega.$$

2.2 Fertility and the distribution of abilities

Let $P_t(\omega)$ be the size of population with ability ω at time $t = 0, 1, 2, ...$, and $P_t \equiv \sum_{\omega \in \Omega} P_t(\omega)$ be total population at time t. The initial distribution of population, $\{P_0(\omega_i)\}_{i=1}^{n}$, is given. Assuming that a law of large number holds, the size of population in a particular income group evolves according to:

$$P_{t+1}(\omega_j) = \sum_{\omega_i \in \Omega} f(\omega_i) P_t(\omega_i) M(\omega_j, \omega_i) \text{ for all } \omega_j \in \Omega. \quad (1)$$

Let $\pi_t(\omega) \equiv P_t(\omega)/P_t$ be the fraction of population with ability $\omega \in \Omega$ at time t. Since income is equal to ability, π also characterizes the income distribution of the economy. The law of motion of π is given by:

$$\pi_{t+1}(\omega_j) = \frac{P_t}{P_{t+1}} \sum_{\omega_i \in \Omega} f(\omega_i) \pi_t(\omega_i) M(\omega_j, \omega_i) \text{ for all } \omega_j \in \Omega \quad (2)$$

Let $\pi^*(\omega) = \lim_{t \to \infty} \pi_t(\omega)$. As shown by CK, the limit is well defined.

A central topic of the paper is to characterize π_t and π^* as well as their relationship to fertility. The following proposition provides a simple but important benchmark. The first part states that when fertility is identical across types the limit distribution of incomes is equal to μ, the invariant distribution associated to M. This result provides a baseline distribution in absence of fertility differences. In that case, the distribution of income just reflects the genetic distribution of abilities, what can be termed nature rather than nurture. The second part of the Proposition shows that fertility differences alone does not necessarily affect the long-run distribution of income, π. In particular, fertility differences are irrelevant for the income distribution when abilities are i.i.d. All proof are in the Appendix.

Proposition 1. When π^* equals μ. Suppose one of the following two assumptions hold: (i) $f(\omega) = f$ for all $\omega \in \Omega$; or (ii) $M(\omega', \omega)$ is independent of ω for all $\omega' \in \Omega$. Then $\pi^*(\omega) = \mu(\omega)$ for all $\omega \in \Omega$.

Fertility differences affect the distribution of incomes when abilities are persistent. The following Proposition is an application of CK's Theorem 2. It states that if the fertility of the poor is higher than the fertility of the rest of the population then π^* is different from μ, and moreover, μ dominates π^* in the first order stochastic sense.

Proposition 2. Suppose M satisfies Assumption 1 and $f(\omega_1) > f(\omega_i) = f$ for all $i > 1$. Then $\sum_{i=1}^{I} \pi^*(\omega_i) > \sum_{i=1}^{I} \mu(\omega_i)$ for all $1 \leq I \leq n$.

Proof. See Chu and Koo (1990, pp.1136).

Comparing Propositions 1 and 2, it follows that a reduction in the fertility of the poor results in a limit distribution that dominates the original distribution. More generally, CK show that if fertility decreases with income and the initial distribution of incomes is at its steady state level, $\pi_0^*(\omega_i)$, then a reduction in the fertility of the poor results in a sequence of income distributions that first order stochastically dominates $\pi_0^*(\omega_i)$, that is, $\sum_{i=1}^{I} \pi_t(\omega_i) < \sum_{i=1}^{I} \pi_0^*(\omega_i)$ for all $1 \leq I \leq n$ and $t > 0$.

2.3 Social Welfare

CK consider average utilitarian welfare functions of the form:

$$\overline{W}(\beta_p) = \sum_{t=0}^{\infty} \sum_{\omega \in \Omega} \beta_p(t) U(\omega) \pi_t(\omega) \tag{3}$$

where $U(\omega)$ is the utility of an individual with ability ω and $\beta_p(t)$ is the weight of generation t in social welfare. A particular case emphasized by CK is one where the planner cares only about steady state welfare: $\beta_p(t) = 0$ for all t and $\lim_{t \to \infty} \beta_p(t) = 1$. In that case,

$$\overline{W}^* = \sum_{\omega \in \Omega} U(\omega) \pi^*(\omega) \tag{4}$$

The following corollary of Proposition 2 provides the theoretical support to family planning programs for the poor, as claimed by CK.

Corollary 3. Suppose social welfare is defined by (3) where $U(\omega)$ is a non-decreasing function of ability. Furthermore, suppose M satisfies Assumption 2 and $f(\omega_1) > f(\omega_i) = f$ for all $i > 1$. Then (i) reducing the fertility of the poor increases social welfare; (ii) fertility policies that do not change the distribution of abilities does not change social welfare.

Corollary 3 holds because reducing fertility of the poor improves the observed distribution of abilities but does not alter $U(\cdot)$. In the next two sections we show counterexamples to Corollary 3 when fertility is endogenous. As a preview of the results, we show a case in which fertility is restricted by policy, the distribution of incomes does not change in any period but social welfare as well as individual welfare decreases for all individuals in all periods compared to the unrestricted case. The reason why the previous Corollary fails to account for this possibility is that it presumes that $U(\omega)$ is invariant to policies, it lacks microfoundations. However, $U(\omega)$ is in fact an indirect utility function and therefore it is not invariant to policies.

3 An Economic Model of Fertility

We now consider the endogenous determination of fertility. Assumptions are the same as in the previous section. In particular, the initial distribution of population across abilities, $\{P_0(\omega_i)\}_{\omega_i \in \Omega}$, is given, abilities are random, determined at birth and described by a Markov chain M satisfying Assumption 1, and having a unique invariant distribution, μ. The technology of production is linear in labor: one unit of labor produces one unit of perishable output. Let $\omega^t = [\omega_0, \omega_1, ..., \omega_t] \in \Omega^{t+1}$ denote a particular realization of ability history up to time t, for a particular family line. There is neither capital nor aggregate risk.

3.1 Individual and aggregate constraints

Markets open every period. The resources of an individual of ability ω_t at time t are labor income and transfers from their parents. Labor income equals $\omega_t (1 - \lambda f_t)$ where λ is the time cost of raising a child. Let $b_t(\omega_t)$ denote transfers, or bequests, received from parents. Resources are used to consume and to leave bequests to children. Insurance market exists as parents can leave bequest contingent on the ability of their children. Let $q_t(\omega^t, \omega_{t+1},)$ be price of an asset that delivers one unit of consumption to a child of ability ω_{t+1} given that the history up to time t is ω^t.[5] The budget constraint of an individual at time t with history ω^t is:

$$c_t\left(\omega^t\right) + f_t\left(\omega^t\right)\sum_{i=1}^{n} q_t(\omega^t, \omega_i) b_{t+1}\left(\omega^t, \omega_i\right) \leq \omega_t\left(1 - \lambda f_t\left(\omega^t\right)\right) + b_t\left(\omega^t\right). \qquad (5)$$

We assume that parents cannot leave negative bequests to their children:

$$b_{t+1}\left(\omega^t, \omega_i\right) \geq 0 \text{ for all } \omega^t \in \Omega^{t+1}, \omega_i \in \Omega \text{ and all } t > 0.$$

Furthermore, suppose $b_0(\omega_i) = 0$ for all $\omega_i \in \Omega$.

Since output is perishable, aggregate consumption must be equal to aggregate production. Alternatively, aggregate savings must be zero. Savings are equal to the total amount of bequests left by parents. Since bequests are non-negative then aggregate savings are zero if and only if all bequests are zero. Therefore, in any equilibrium the budget constraints (5) simplifies to:

$$c_t\left(\omega^t\right) \leq \omega_t\left(1 - \lambda f_t\left(\omega^t\right)\right) \text{ for all } \omega^t \in \Omega^{t+1} \text{ and all } t \geq 0. \qquad (6)$$

This is balanced budget constraint for every period and state. The lack of intergenerational transfers significantly simplifies the problem and explain why social mobility arises in the equilibrium. Otherwise, as shown by Alvarez (1999), parents will use family size to buffer against shocks and use transfers to smooth consumption across time and states regardless of ability preventing thus any social mobility. Absent transfers, ability becomes the key determinant of consumption and fertility, as we see below.

In addition to budget constraints, individuals must satisfy time constraints. In particular, the time spent in raising children cannot exceed the time available to an individual, which is

[5] The price also depends on the aggregate distribution of abilities at time t.

normalized to 1. Thus,

$$0 \leq f_t\left(\omega^t\right) \leq \frac{1}{\lambda}. \tag{7}$$

3.2 Individual's Problem

The lifetime utility of an individual born at time t is of the Barro-Becker type (Barro and Becker 1989 and Becker and Barro 1988):

$$U_t = u\left(c_t\right) + \beta f_t^{1-\epsilon} E_t U_{t+1},\ t = 0, 1, 2, ... \tag{8}$$

where $u\left(c\right) = \frac{c^\sigma}{\sigma}$, $\sigma \in (0, 1)$, is the utility from consumption, f_t is the number of children, U_{t+1} is the utility of the time $t+1$ generation, and E_t is the mathematical expectation operator conditional on the information up to time t. The term $\beta f_t^{1-\epsilon}$ is the weight that parents place on their f_t children. When $\epsilon = 0$ parents are perfectly altruistic toward children. We assume $0 \leq \epsilon < 1$.

The following restrictions on parameters are needed in order to have a well-behaved bounded problem.

Assumption 3. $1 - \epsilon > \sigma$ and $\lambda^{1-\epsilon} > \beta$.

The first part of the assumption is identical to the one discussed by Barro and Becker (1988) to assure strict concavity of the problem. The second part guarantees bounded utility as the effective discount factor in that case satisfies $\beta f_t^{1-\epsilon} \leq \beta\lambda^{\epsilon-1} < 1$.[6]

The individual's problem is to choose a sequence $\left\{f_t\left(\omega^t\right)\right\}_{t=0}^{\infty}$ to maximize U_0 subject to (6) and (7). The problem can be written in sequence form, by recursively using (8), to obtain:

$$U_0^*\left(\omega_0\right) = \max_{\left\{P_{t+1}\left(\omega^{t-1},\omega_t\right)\right\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t P_t\left(\omega^{t-1}\right)^{1-\varepsilon} u\left(\omega_t\left(1 - \lambda\frac{P_{t+1}\left(\omega^{t-1},\omega_t\right)}{P_t\left(\omega^{t-1}\right)}\right)\right) \tag{9}$$

[6] An upper bound for U_t is $\frac{u(\omega_n)}{1-\beta\lambda^{\epsilon-1}}$.

subject to

$$0 \leq P_{t+1}\left(\omega^{t-1}, \omega_t\right) \leq P_t\left(\omega^{t-1}\right)/\lambda \text{ for all } \omega^{t-1} \in \Omega^t,\ \omega_t \in \Omega \text{ and } t \geq 0,\ \ P_0 > 0.$$

In this formulation, $P_0\left(\omega^{-1}\right) = 1$. $P_{t+1}\left(\omega^t\right) = \prod_{j=0}^{t} f_j\left(\omega^j\right)$. Fertility rates can be recovered as $f_t\left(\omega^t\right) = \frac{P_{t+1}\left(\omega^{t-1},\omega_t\right)}{P_t\left(\omega^{t-1}\right)}$.

An alternative way to describe the household problem is by the following functional equation:

$$U(\omega) = \max_{f \in [0, \frac{1}{\lambda}]} u\left(\omega\left(1 - \lambda f\right)\right) + \beta f^{1-\epsilon} E\left[U\left(\omega'\right)|\omega\right] \tag{10}$$

The next proposition states that the principle of optimality holds for this problem. This result is novel because the functional equation is not standard due to the endogeneity of fertility. In particular the discount factor is endogenous. Alvarez (1999) shows that the principle of optimality holds for a dynastic version of this problem, while we show that it holds for the household version of the problem.[7] Our household problem is simpler because of the lack of intergenerational transfers in equilibrium.

Proposition 4. The functional equation (10) has a unique solution, $U(\omega)$. Moreover $U(\omega) = U_0^*(\omega)$ for $\omega \in \Omega$.

[7] The analogous dynastic problem is:

$$V(N, \omega) = \max_{N' \in [0, \frac{1}{\lambda} N]} u\left(\omega - \lambda N'/N\right) N^{1-\epsilon} + \beta E\left[U\left(N', \omega'\right)|\omega\right].$$

In this problem the number of family members is a state variable, N, all member have the same ability, ω, and make the same choices. The household problem does not impose these constraints.

3.3 Optimal Fertility

The optimality condition for an interior fertility choice is:[8]

$$\lambda\omega u'\left(\left(1-\lambda f^{*}\right)\omega\right)=\beta\left(1-\epsilon\right)f^{*-\epsilon}E\left[U\left(\omega'\right)|\omega\right] \tag{11}$$

Let $f^{*}=f(\omega)$ be the optimal fertility rule and $c^{*}=c\left(\omega\right)\equiv\left(1-\lambda f(\omega)\right)\omega$ the optimal consumption rule. The left hand side of this expression is the marginal cost of a child while the right hand side is the marginal benefit. The marginal cost is the product of the cost per child, $\lambda\omega$, times the marginal utility of consumption. The marginal benefit to the parent is the expected welfare of a child, $E\left[U\left(\omega'\right)|\omega\right]$, times the parental weight associated to the last child, $\beta\left(1-\epsilon\right)f(\omega)^{-\epsilon}$. According to (11), both the marginal cost and the marginal benefit of children are increasing functions of ability, ω. The marginal cost increases with ω because it increases the opportunity cost of the parental time required to raise children. The marginal benefit increases because of the postulated intergenerational persistence of abilities: high ability parents are more likely to have high ability children.

It is instructive to write the first order condition in an alternative way. First, use equation (11) to express (10) as:

$$U_{t}=u\left(c\left(\omega\right)\right)+\frac{1}{1-\epsilon}f\left(\omega\right)\lambda\omega u'\left(c\left(\omega\right)\right) \tag{12}$$

Then use (12) to rewrite (11) as:

$$u'\left(c\left(\omega\right)\right)\left(f\left(\omega\right)\right)^{\epsilon}=\beta E\left[u'\left(c\left(\omega'\right)\right)\frac{\omega'}{\omega}\left(\frac{1}{\lambda}+\frac{1-\epsilon-\sigma}{\sigma}\left(\frac{1}{\lambda}-f\left(\omega'\right)\right)\right)\Big|\,\omega\right]. \tag{13}$$

This equation is useful because it only requires marginal utilities, rather than total utility as in equation (11), and corresponds to the Euler Equation of the problem describing the optimal

[8] Corner solutions are not optimal due to the properties of utility functions and altruistic function. Having no children is never optimal because the marginal benefit of a child is infinite while the marginal cost is finite. In particular, notice that $E\left[U\left(\omega'\right)|\omega\right]>0$ for all ω while $\lim_{f\to 0}f^{-\epsilon}=\infty$. Having the maximum number of children is also sub-optimal because the marginal cost is infinite when parental consumption is zero.

consumption rule. Although savings are zero in equilibrium, fertility allows individuals to smooth consumption across generations.[9]

To better understand the implications of the model it is useful to consider some specific cases. The following Proposition consider three cases: i.i.d abilities across generations, perfect intergenerational persistence of abilities with no uncertainty and random walk (log) abilities[10]

Proposition 5. Persistence and the fertility-ability relationship. (i) Fertility decreases with ability if abilities are i.i.d across generations. In this case $f(\omega)$ satisfies the equation $\frac{f(\omega)^{\epsilon}}{(1-\lambda f(\omega))^{1-\sigma}} = A\omega^{-\sigma}$ where A is a constant. Furthermore, fertility is independent of ability in one of the following two cases: (ii) M is the identity matrix (abilities are perfectly persistent and deterministic); or (iii) $\ln \omega_t = \ln \omega_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim N(0, \sigma^2)$.

According to Proposition 5, fertility decreases with ability when abilities are i.i.d. The intuition is that without intergenerational persistence, the ability of the parent only affects her/his marginal cost but not her/his marginal benefit as $E[U(\omega')|\omega] = E[U(\omega')]$ for all $\omega \in \Omega$. On the other extreme, fertility is independent of ability when abilities are perfectly persistent across generations (cases ii and iii). This is because in those cases both the marginal cost and the marginal benefit are proportional to ω^{σ}.

Given that fertility becomes only independent of ability in the extreme case of perfect persistent, it is natural to conjecture that fertility decreases with ability when persistence is less than perfect. We were able to confirm this conjecture numerically but analytical solutions were not obtained.

[9] Equation (11) can also be written in the form of a more traditional Euler Equation. Let $1 + r'$ be the gross return of "investing" in a child. It is given by $1 + r' \equiv \frac{U(\omega')/u'(c')}{\lambda\omega}$. In this expression, $U(\omega')/u'(c')$ is the value of a new life, in terms of goods, while $\lambda\omega$ is the cost of creating a new individual. Then (11) can be written as:

$$u'(c) = \beta(1-\epsilon) f^{*-\epsilon} E\left[u'(c')\left(1+r'\right)|\omega\right]. \tag{14}$$

This is an Euler Equation with a discount factor $\beta(1-\epsilon) f^{*-\epsilon}$. It suggests that optimal fertility choices are similar to saving decisions and that children are like an asset, as pointed out by Alvarez (1999). However, two important differences with the traditional Euler Equation are that the individual controls both the discount factor and the gross return.

[10] Although a random walk does not satisfy some of the assumptions above, it helps to develop some intuition.

3.4 Dynamics of the Income Distribution

Given the optimal fertility rule $f(\cdot)$, initial distribution $\pi_0(\cdot)$ of population across abilities, and initial population P_0, distributions of income for all periods can be obtained using equations (1) and (2). Furthermore, average earning abilities and average income are given by:

$$E_t = \sum_{\omega\in\Omega} \omega\pi_t(\omega);\ I_t = \sum_{\omega\in\Omega} \omega(1-\lambda f(\omega))\pi_t(\omega)$$

In the next section we use the microfounded model to perform welfare evaluations of family planing programs. The model also allows us to assess whether Assumption 1 and Assumption 2 are somewhat associated. They are. A mobility matrix with less than perfect persistence of intergenerational abilities can give rise to a negative relationship between fertility and ability. The following Proposition revisits Proposition 1 at the light of the micro-founded model. It plays an important role in section 4 when providing counter-examples to CK's claims.

Proposition 6. Persistence, fertility and ability distribution. (i) If $M(\cdot,\omega)$ is independent of ω then $f(\omega)$ decreases with ω and $\pi^*(\omega) = \mu(\omega)$ for $\omega \in \Omega$; (ii) if M is the identity matrix then $f(\omega) = f$ for all $\omega \in \Omega$ and $\pi_t(\omega) = \pi^*(\omega) = \pi_0(\omega)$ for all $\omega \in \Omega$ and all t; (iii) if $\ln\omega$ follows a Gaussian random walk then $f(\omega) = f$, and given ω_0 the variance of abilities diverges to ∞.

In words, if abilities are i.i.d. across generations, then fertility decreases with ability but the observed limit distribution of abilities is independent of fertility choices and equal to $\mu(\omega)$. Furthermore, with certainty and perfect intergenerational persistence of abilities the observed distribution of abilities in any period is identical to the initial distribution of abilities. Finally, if (log) abilities follow a random walk then there is not limit distribution of abilities since its variance goes to infinite.

3.5 Fertility Policies and Individual Welfare

Consider now a family planning policy that sets lower and/or upper bounds on fertility choices. Let $\underline{f}(\omega) \geq 0$ and $\overline{f}(\omega) \leq 1/\lambda$ be the lower and upper bound respectively. Bounds potentially depend on individual abilities. The indirect utility $U^r(\omega)$ of the constrained problem is described by the following Bellman equation:

$$U^r(\omega) = \max_{f \in [\underline{f}(\omega), \overline{f}(\omega)]} u((1-\lambda f)\omega) + \beta f^{1-\epsilon} E\left[U^r(\omega') \,|\omega\right]. \tag{15}$$

Let $f^r(\omega)$ denotes the optimal fertility rule. The following Proposition is one of the main results of the paper. It states that binding fertility restrictions in at least one state reduces the indirect utility, or welfare, of all individuals even those whose fertility is not directly affected. The Proposition also states that fertility restrictions of any type (weakly) reduces the fertility of all individuals except perhaps those whose fertility rates are at or below the lower bound.

Proposition 7. $U^r(\omega) \leq U(\omega)$ *with strict inequality if* $f(\omega) > \overline{f}(\omega)$ *or* $f(\omega) < \underline{f}(\omega)$ *for at least one* ω*. Furthermore,* $f^r(\omega) = \underline{f}(\omega)$ *if* $f(\omega) \leq \underline{f}(\omega)$ *and* $f^r(\omega) \leq f(\omega)$ *otherwise.*

Fertility restrictions reduce welfare because it restricts individuals' choices without providing any compensation. Furthermore, fertility restrictions that only affects a particular group, say the lowest ability individuals, results in lower welfare for all individuals because, regardless of current ability, there is a positive probability that a descendant of the dynasty will fall into the group directly affected in finite time. Proposition 7 implies that policies such as the *One Child Policy*, which imposes a uniform bound on all ability levels, or policies that compel individuals to increase their fertility, such as the "leftover" women stigma in China, are detrimental to individual welfare, according to our model. Given that welfare of all individuals falls, the marginal benefits of having children also falls while the marginal cost remains the same. As a result, fertility must fall for all types except perhaps for those who are constrained by the policy to increase their fertility. We next study the consequences of fertility restrictions on social welfare.

4 Family Planning and Social Welfare Reconsidered

Given that fertility policies reduces the welfare of all individuals, as stated in Proposition 7, it is natural to infer that social welfare should also fall. The answer, however, depends on how social welfare is defined and whether the policy reduces or increases population. In this section we focus on fertility policies that impose upper limits on fertility rates such as limiting the fertility of the poor or the *One Child Policy*.

4.1 Analytical results

According to Proposition 7, upper limits on the fertility of any ability group reduces fertility of all ability groups. Therefore, upper limits on fertility unequivocally reduces population of all ability groups at all times after time 0. Given that both population and individual welfare fall for all ability types, we are able to show that fertility limits unequivocally decrease social welfare if social welfare is of the classical, or Bentham, utilitarian form. Classical utilitarianism defines social welfare as the total discounted welfare of all (born) individuals:[11]

$$W = \sum_{t=0}^{\infty} \sum_{\omega \in \Omega} \beta_p(t) U(\omega) P_t(\omega) . \tag{16}$$

In this formulation $\beta_p(t) \geq 0$ is the weight the social planner assigns to generation t. Since individuals are altruistic toward their descendants, $\beta_p(t) > 0$ means that the planner gives additional weight to generation t on top of what is implied by parental altruism. A particular case in which the planner weights only the original generation, and therefore adopts its altruistic weights, is the one with $\beta_p(0) = 1$ and $\beta_p(t) = 0$ for $t > 0$:

$$W_0 = \sum_{\omega \in \Omega} U(\omega) P_0(\omega) \tag{17}$$

[11] The results are similar if the welfare of the unborn is explicitly considered as long as the unborn enjoy lower utility than the born.

The following Proposition states the main conclusion of the paper: restricting fertility decreases classical utilitarian welfare.

Proposition 8. Imposing upper limits on fertility choices reduces social welfare as defined by (16).

An identical result is obtained if the planner exhibits positive but diminishing returns to population, say if $P_t(\omega)$ in expression (16) is replaced by $P_t(\omega)^{1-\epsilon_p}$ where $\epsilon_p \in (0,1)$. This formulation seems the natural extension of the Barro-Becker preferences for a planer.

An alternative definition of social welfare is average, or Mills, utilitarianism as defined by equation (3) for the general case, and with (4) as a special case. This definition of welfare, the one used by CK, is analogous to (16) but uses population shares, $\pi_t(\omega)$, rather than population, $P_t(\omega)$. Under this definition, social welfare could increase even if the welfare of all individuals fall. The net effect depends on the relative strength of two potentially opposite forces: on the one hand individual welfare falls but on the other hand the distribution of abilities, π, may improve, as in CK. We next show analytical examples in which social welfare falls even under this definition. These are analytical formal counterexamples to CK's claims that fertility limits on the poor are welfare enhancing.

The following Proposition states that fertility restrictions of any type reduce average utilitarian welfare if abilities are i.i.d.

Proposition 9. Suppose $M(\cdot, \omega)$ is independent of ω for all $\omega \in \Omega$ and $\pi_0(\omega_i) = \mu(\omega_i)$. Then upper limits on fertility choices reduce social welfare as defined by (3).

Proposition 9 relies on the earlier finding in Proposition 1 that, when abilities are i.i.d, the distribution of abilities among the population is independent of fertility choices and the limit distribution of ability is the invariant distribution of M. We show in the appendix that $\pi_t(\omega_i) = \mu(\omega_i)$ for all t if $\pi_0(\omega_i) = \mu(\omega_i)$. Therefore, in the i.i.d case the effect of any fertility policy on social welfare, as defined by (3), is only determined by its effect on individual welfare, U.

A particular implication of Proposition 9 is that limiting the fertility of the poor reduces welfare which contradicts CK's claim stated in Corollary 3. The i.i.d case in Proposition 9 satisfies CK's Assumptions 1 and 2 since fertility rates are decreasing, as stated in Proposition 5, and i.i.d abilities satisfies conditional stochastic monotonicity. Corollary 3 fails to properly describe the effect of the policy on social welfare because it implicitly assumes that U is unaffected by the policy change.

The following is a deterministic example showing that average utilitarian welfare unequivocally falls with "uniform" fertility restrictions such as the one child policy.

Proposition 10. Suppose M is the identity matrix and $\overline{f}(\omega) = \overline{f}$. Then fertility restrictions reduces social welfare as defined by (3).

Proposition 10 provides another example in which fertility restriction do not affect π. Since in the deterministic case all ability groups have the same fertility choices, and the fertility restriction affect all ability groups equally, then it follows that $\pi_t = \pi_0$ for all t so that the effect of the policy on social welfare is only determined by the effect on individual welfare U.

We now turn to numerical simulations to investigate more generally the effects of fertility policies on social welfare.

4.2 Calibration and Simulations

4.2.1 Benchmark calibration

The following parameters are needed to simulate the model: the Markov process of abilities M, preference parameter σ, altruistic parameters β and ϵ, cost of raising children λ, and social planner weight $\beta_p(t)$.

Income groups, fertilities of different income groups, and the Markov chain are taken from Lam (1986) who provides estimates for Brazil. Average incomes for five income groups are $\overrightarrow{I} = [553, 968, 1640, 2945, 10991]$. They describe income classes of Brazilian male household heads aged from 40 to 45 in 1976. Average fertility of each income group are $\overrightarrow{f} =$

$[6.189, 5.647, 5.065, 4.441, 3.449]/2$. We divide fertility by two to obtain fertility per-adult. Using income and fertility data, we calculate earning abilities of different groups as $\omega_i = \frac{I(\omega_i)}{1-\lambda f(\omega_i)}$ and normalize the lowest ability to be 1. The Markov chain provided by Lam is:

$$M = \begin{bmatrix} 0.50 & 0.25 & 0.15 & 0.10 & 0.05 \\ 0.25 & 0.40 & 0.20 & 0.20 & 0.10 \\ 0.15 & 0.20 & 0.35 & 0.20 & 0.20 \\ 0.05 & 0.10 & 0.20 & 0.35 & 0.25 \\ 0.05 & 0.05 & 0.10 & 0.15 & 0.40 \end{bmatrix}$$

This chain does not satisfy conditional stochastic monotonicity property although its diagonal elements dominate other elements implying certain level of earning persistency across generations. We also consider the Markov chain provided by CK, which satisfies CSM, and obtain similar results. Initial population is normalized to 1. The initial distribution of abilities, π_0, is approximated by the stationary distribution implied by M and $\overrightarrow{f}$.

Our altruistic function, $\beta f^{1-\epsilon}$, is calibrated following Manuelli and Seshadri (2009) (MS henceforth).[12] For σ we initially used MS's parameter of 0.38. However, the fertility rates implied by the calibrated model were too high and the range of fertilities too small compared to Brazilian fertility data. We set $\sigma = 0.68$ to better fit the fertility data. Another key parameter of the model is the time cost of raising a child, λ. We choose $\lambda = 0.2$ which implies a maximum number of 10 children per couple, or that each parent spend 10% of their time on every child. We perform robustness checks for this and other parameters. For the social planner weights we assume $\beta_p(t) = \delta^t$ with $\delta = 0.1$. The set of parameters used for the benchmark exercises are summarized in Table 1.

[12] Their altruistic function takes the form $e^{-\rho B} e^{-\alpha_0 + \alpha_1 \ln f}$ where $B = 25$ is the age of fertility. So the proper mapping is $\beta = e^{-\rho B} e^{-\alpha_0}$ and $1 - \epsilon = \alpha_1$.

Table 1. Parameters Setting

Parameters	Concept	Values
β	individual discount factor	0.29
σ	elasticity of substitution	0.68
ϵ	altruistic parameter	0.35
λ	per child time cost	0.2
δ	weight of social planner	0.1

4.2.2 Results

The simulated model reproduces a negative relationship between fertility and ability similar to the Brazilian data.[13] Because abilities are persistent but not perfectly persistent across generations, the increase of the marginal cost of children dominates that of the marginal benefit as ability increases. As shown in the first panel of Figure 1, fertility per household falls from 9 to 2 as earning abilities increase from 1 to 12. The second panel plots average ability, $\overline{\omega}$, and average income, $\overline{y}$, as the upper bound of fertility increases. As predicted by CK, tighter fertility limits, which affect lower income groups more severely, increase average income and ability.

The remaining panels in Figure 1 illustrate the effect of fertility limits on various welfare measures. On the horizontal axis is the uniform fertility upper limit imposed on all ability groups, a limit that goes from 0 to 10 children per household. It shows that steady state average welfare, $\overline{W}^*$, average welfare of all generations, $\overline{W}(\beta_p)$, welfare of the initial generation, W_0, and total welfare of all generations, $W(\beta_p)$, all increase as the upper bound on fertility is relaxed. These results confirm the main message of the paper: fertility restrictions, on the poor or other groups, do not have strong theoretical support for improving people's welfare.

We also study the welfare effects of imposing lower bounds on fertility rates. This type of restrictions disproportionately affect the rich, or high ability individuals, because their unconstrained fertility is typically lower. Figure 2 shows that this policy increases average ability since high ability individuals have proportionally more high ability children. On the other hand, the policy reduces average income because individuals, especially those with high

[13] By construction, our calibration targets the dispersion of fertilities but not the sign of the relationship between fertility and income.

ability, spend more time raising children and this effect dominates the effect of an improved ability distribution. All four welfare measures unanimously decrease as the lower bound on fertility increases.

In summary, the results above show that fertility restrictions, on the poor and on the rich, does not result into higher social welfare although they may improve the distribution of abilities and income.

4.2.3 Robustness Checks

We now report the results of various robustness checks. For this purpose we change one parameter at a time while keeping all the other parameters at their benchmark values and study the effect on the various welfare measures of imposing an upper limit on fertility. We find that the qualitative results obtained above are mostly robust although there exists a set of parameters for which average steady state welfare, $\overline{W}^*$, improves with fertility restrictions. The set of parameters studied is further restricted by the need to have finite utility.

The results are robust to setting σ below 0.74. If σ is larger than 0.74, relaxing fertility restrictions slightly reduce steady state average welfare, $\overline{W}^*$, but only when there is a tight upper limit on fertility, of between 1 and 2, as illustrated in the first panel of Figure 3 for $\sigma = 0.9$. A high elasticity of intergenerational substitution significantly reduces the gains of smoothing consumption through fertility choices. In addition, low fertility allows higher consumption and low marginal utility of consumption. As a result, a relaxation of fertility restrictions have a minor impact on individual utility and the change in the distribution of abilities determines the change in social welfare. However, further relaxation of the upper limit increases $\overline{W}^*$.

We also find that if β is sufficiently low, a tighter fertility restriction may increase $\overline{W}^*$ as illustrated in the second panel of Figure 3 for the case $\beta = 0.2$. A low β means that parents care little about future generations, have fewer kids, higher consumption and lower marginal utility of consumption. In this case, fertility restrictions have a minor effect on individual welfare and, as a result, the change of the ability distribution is the dominant effect determining social welfare. However, this low degree of altruism also implies that the model predicts counterfactually low fertility rates. A similar result is obtained when ϵ is particularly large, as illustrated in the third panel of Figure 3 for $\epsilon = 0.53$.

Finally, if the cost of raising children, λ, is sufficiently large then a tighter fertility restriction may increase $\overline{W}^*$ as shown in the last panel of Figure 3 for the case $\lambda = 0.28$. In this case the high cost of raising children itself prevents households from having many children and therefore fertility restrictions are not very harmful for individual welfare. The change in social welfare is therefore primarily determined by the change in the distribution of abilities.

5 Conclusion

Stochastic dominance, or welfare dominance, seemingly provides a robust criterion for policy evaluation. It allows to rank policies by simply looking at the resulting income distribution without requiring much knowledge of individuals' preferences and constraints, or knowledge of the social welfare function. Cho and Koo (1990) exploit such apparent generality to provide a striking policy recommendation. They assert that stochastic dominance "provides us with very strong theoretical support in favor of family-planning programs that encourage the poor in developing countries to reduce their reproductive rate (pp 1136)." Such fundamental claim has surprisingly remained unchallenged. In this paper we show that stochastic dominance alone does not provide the strong theoretical support claimed by CK. Our findings challenges not only CK's main normative conclusion but also the larger classical literature on the topic of welfare dominance which is the foundation of such conclusion.

Our main contribution is to provide explicit micro-foundations to CK's model. The key features are altruism, random abilities, labor costs of raising children, non-negative bequest constraints, and an endowment economy. The model is particularly useful because its equilibrium exactly maps into the Markov branching framework of CK. It also successfully replicates two basic features of the evidence on fertility and income distribution: fertility decreases with ability and social mobility occurs in equilibrium. These features are not easily obtained by altruistic models of fertility.

We show that fertility restrictions reduces social welfare in our model in spite of the fact that they may result in superior income distributions in the first order stochastic sense. Contrary to CK, and to a larger literature mentioned in the introduction, we find that first order stochastic dominance does not provide a strong theoretical support to family-planning programs directed toward reducing the fertility of the poor. The main reason for this failure is that stochastic dominance does not account for the fact that indirect utility functions are not invariant to fertility policies.

Our model abstracts from a number of aspects that are potentially important to fertility decisions such as bequests, human capital accumulation, and wealth inequality. We study these extensions in Cordoba el al. (2013). The models are significantly more complicated, and do not map into a simple Markov branching framework, but our early results confirm the findings that policies restricting fertility typically does not increase social welfare.

References

[1] Alvarez, F. (1999), "Social Mobility: The Barro-Becker Children Meet the Laitner-Loury Dynasties", *Review of Economic Dynamics*, Vol. 2, pp. 65-103.

[2] Atkinson, A. B. (1970), "On the Measurement of Inequality", *Journal of Economic Theory*, Vol. 2, pp. 244-263.

[3] Atkinson, A. B. and Brandolini, A. (2010), "On Analyzing the World Distribution of Income", *World Bank Economic Review*, Vol. 24, pp. 1-37.

[4] Barro, R. J. and Becker, G. S. (1989), "Fertility Choice in a Model of Economic Growth", *Econometrica*, Vol. 57, pp. 481-501.

[5] Becker, G. S. (1960), "An Economic Analysis of Fertility, in Demographic and Economic Change in Developed Countries", *Conference of the Universities, National Bureau Committee for Economic Research*, Princeton University Press, pp. 209-240.

[6] Becker, G. S. (1983), "Some Notes on Population Growth and Economic Growth", *Unpublished manuscript.*

[7] Becker, G. S. and Barro, R. J. (1988), "A Reformulation of the Economic Theory of Fertility", *The Quarterly Journal of Economics*, Vol. 103, pp. 1-25.

[8] Chu, C. Y. C. and Koo, H.-W. (1990), "Intergenerational Income-Group Mobility and Differential Fertility", *American Economic Revie*w, Vol. 80, pp. 1125-1138.

[9] Cordoba, J. C., Liu, X. and Ripoll, M. (2013), "Endogenous Fertility and the Distribution of Wealth", *Mimeo.*

[10] Cordoba, J. C. and Ripoll, M. (2010), "A Contribution to the Economic Theory of Fertility", *Mimeo.*

[11] Dasgupta, P., Sen, A. and Starrett, D. (1973), "Notes on the Measurement of Inequality", *Journal of Economic Theory*, Vol. 6, pp. 180-187.

[12] Davidson, R. and Duclos, J.-Y. (2000), "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality", *Econometrica*, Vol. 68, pp. 1435-1464.

[13] Eckstein, Z. and Wolpin, K. I. (1985), "Endogenous Fertility and Optimal Population Size", *Journal of Public Economics*, Vol. 27, pp. 93-106.

[14] Foster, J. E. and Shorrocks, A. F. (1988a), "Poverty Orderings", *Econometrica*, Vol. 56, pp. 173-77.

[15] Foster, J. E. and Shorrocks, A. F. (1988b), "Poverty Orderings and Welfare Dominance", *Social Choice and Welfare*, Vol. 5, pp. 179-198.

[16] Golosov, M., Jones, L. E. and Tertilt, M. (2007), "Efficiency with Endogenous Population Growth", *Econometrica*, Vol. 75, pp. 1039-1071.

[17] Hansen, G. D. and Prescott, E. C. (2002), "Malthus to Solow", *American Economic Review*, Vol. 92 (4), pp. 1205-1217.

[18] Jones, L. E. and Tertilt, M. (2006), "An Economic History of Fertility in the U.S.: 1826-1960", *In Frontiers of Family Economics*, Vol. 1, P. Rupert (eds.), Bingley, UK: Emerald Press, pp. 165– 230.

[19] Kolm, S. (1969), "The Optimal Production of Justice", *Public Economics*, pp. 145-200.

[20] Kremer, M. (1993), "Population Growth and Technological Change: One Million B.C. to 1990", *Quarterly Journal of Economics*, Vol. 108, pp. 681-716.

[21] Lam, D. A. (1986), "The Dynamics of Population Growth, Differential Fertility, and Inequality", *American Economic Review*, Vol. 76, pp. 1103-1116.

[22] Lam, D. A. (1993), "Demographic Variables and Income Inequality", in M. R. Rosenzweig and O. Stark, eds, *Handbook of Population and Family Economics*, Vol. 1, Part B, Elsevier, Chapter 18, pp. 1015-1059.

[23] Liao, P.-J. (2013), “The One-Child Policy: A Macroeconomic Analysis", *Journal of Development Economics*, Vol. 101, pp. 49-62.

[24] Lucas, R. Jr. (1976), "Econometric Policy Evaluation: A Critique", Carnegie-Rochester Conference Series on Public Policy, Vol. 1(1), pp. 19-46.

[25] Manuelli, R. E. and Seshadri, A. (2009), “Explaining International Fertility Differences", *The Quarterly Journal of Economics*, Vol. 124, pp. 771-807.

[26] Pazner, E. A. and Razin, A. (1980), “Competitive Efficiency in an Overlapping-generation Model with Endogenous Population", *Journal of Public Economics*, Vol. 13, pp. 249-258.

[27] Raut, L. K. (1990), “Capital Accumulation, Income Distribution and Endogenous Fertility in an Overlapping Generations General Equilibrium Model", *Journal of Development Economics*, Vol. 34, pp. 123-150.

[28] Rothschild, M. and Stiglitz, J. E. (1973), “Some Further Results on the Measurement of Inequality", *Journal of Economic Theory*, Vol. 6, pp. 188-204.

[29] Saposnik, R. (1981), “Rank-dominance in Income Distributions", *Public Choice*, Vol. 36, pp. 147-151.

[30] Saposnik, R. (1983), “Ranking Income Distributions", *Econometrica*, Vol. 50, pp. 3-17.

[31] Willis, R. J. (1985), “Externalities and Population", in D. Gale Johnson and Ronald D. Lee (eds.), *Population Growth and Economic Development: Issues and Evidence*, Madison, WI: University of Wisconsin Press.

Appendix

Proof of Proposition 1 (i) If fertility is exogenously the same for every individual, divide both sides of (1) by P_{t+1}.

$$\frac{P_{t+1}(\omega_j)}{P_{t+1}} = \frac{P_t}{P_{t+1}} \sum_{\omega_i \in \Omega} f(\omega_i) \frac{P_t(\omega_i)}{P_t} M(\omega_j, \omega_i) \text{ for all } \omega_j \in \Omega$$

Using the definition of π_t,

$$\begin{aligned}\pi_{t+1}(\omega_j) &= \frac{P_t}{P_{t+1}} \sum_{\omega_i \in \Omega} f(\omega_i)\, \pi_t(\omega_i)\, M(\omega_j, \omega_i) \\ &= \frac{fP_t}{P_{t+1}} \sum_{\omega_i \in \Omega} \pi_t(\omega_i)\, M(\omega_j, \omega_i) = \sum_{\omega_i \in \Omega} \pi_t(\omega_i)\, M(\omega_j, \omega_i)\end{aligned}$$

The last equality holds because

$$P_{t+1} = P_t \sum_{\omega_i \in \Omega} f\pi_t(\omega_i) = P_t f$$

Taking limit to both sides of the expression with π, we get

$$\pi^*(\omega_j) = \lim_{t\to\infty} \pi_{t+1}(\omega_j) = \lim_{t\to\infty} \sum_{\omega_i \in \Omega} \pi_t(\omega_i)\, M(\omega_j, \omega_i) = \sum_{\omega_i \in \Omega} \pi^*(\omega_i)\, M(\omega_j, \omega_i)$$

Hence $\pi^*(\cdot) = \mu(\cdot)$ is the invariant distribution of M.

Proof of Proposition 1 (ii) $M(\cdot\,, \omega_i)$ is independent of ω_i implies $M(\omega_j, \omega_i) = M(\omega_j)$ for every $\omega_j \in \Omega$. By (1),

$$\begin{aligned}\pi_{t+1}(\omega_j) &= \frac{M(\omega_j)}{P_{t+1}} \sum_{\omega_i \in \Omega} f(\omega_i)\, \pi_t(\omega_i)\, P_t \\ &= M(\omega_j) = \sum_{\omega_i \in \Omega} \pi_t(\omega_i)\, M(\omega_j)\end{aligned}$$

Taking limit to both sides, $\pi^* (\cdot) = \underset{t\rightarrow\infty}{lim} \pi_t (\cdot)$ is equal to the invariant distribution.

Proof of Proposition 4 We first show that there exists a solution $U(\cdot)$ that solves the functional equation (10). Define a set of functions.

$$S = \{ f : \Omega \rightarrow \mathbb{R} | \, \|f\| \leq M \}$$

where $M = \frac{u(\omega_n)}{1-\beta\lambda^{\epsilon-1}}$, and $\|\cdot\|$ is the sup norm. We can show that S is a complete metric space. Define operator T as

$$TU(\omega) = \max_{0\leq f \leq \frac{1}{\lambda}} u((1-\lambda f)\omega) + \beta f^{1-\varepsilon} E\left[U(\omega')\,|\omega\right] \tag{18}$$

for all $\omega \in \Omega$ and $U \in S$. First show that T is a contraction. It suffices to show that T satisfies two properties, monotonicity and discounting. Suppose the sequential problem has a unique solution, then the right hand side of (18) has a solution. Standard argument can show that for any U and $\tilde{U} \in S$ satisfying $U(\omega) \leq \tilde{U}(\omega)$ for all ω, then $TU(\omega) \leq T\tilde{U}(\omega)$ for all ω. The following arguments show discounting property holds. For any given constant b,

$$\begin{aligned} T(U(\omega)+b) &= \max_{0\leq f \leq \frac{1}{\lambda}} u((1-\lambda f)\omega) + \beta f^{1-\varepsilon} E\left[U(\omega') + b|\omega\right] \\ &= \max_{0\leq f \leq \frac{1}{\lambda}} u((1-\lambda f)\omega) + \beta f^{1-\varepsilon} E\left[U(\omega')\,|\omega\right] + \beta b f^{1-\varepsilon} \\ &\leq \max_{0\leq f \leq \frac{1}{\lambda}} u((1-\lambda f)\omega) + \beta f^{1-\varepsilon} E\left[U(\omega')\,|\omega\right] + \beta b \left(\frac{1}{\lambda}\right)^{1-\varepsilon} \\ &= TU(\omega) + \beta\lambda^{\epsilon-1} b \end{aligned}$$

By Contraction Mapping Theorem, there exists a unique fixed point $U : \Omega \rightarrow \mathbb{R}$ that solves the functional equation $TU = U$. The functional equation has a unique solution. The existence of a solution $U(\cdot)$ has been proved, we next show $U(\omega) = U_0^*(\omega)$ for all $\omega \in \Omega$. $\left\{P_{t+1}(\omega^t)\right\}_{t=0}^{\infty}$ are the choice variables in the sequential problem (9). Since the current population $P_t\left(\omega^{t-1}\right)$ is given, the problem is the same if we choose $\left\{f_t\left(\omega^t\right)\right\}_{t=0}^{\infty}$ instead of $\left\{P_{t+1}\left(\omega^t\right)\right\}_{t=0}^{\infty}$, and it can be written as follows. Given ω_i, the welfare of an individual in generation i is

$$U_i^*\left(\omega_i\right)=\max_{\left\{f_t\left(\omega^t\right)\right\}_{t=0}^{\infty}} E_0\sum_{t=i}^{\infty}\beta^{t-i}\prod_{j=i}^{t-1}f_j\left(\omega^j\right)^{1-\epsilon}u\left(\omega_t\left(1-\lambda f_t\left(\omega^t\right)\right)\right)$$

subject to $0\leq f_t\left(\omega^t\right)\leq\frac{1}{\lambda}$. Since at optimal there is no intergenerational transfer of wealth from parents to children, the welfare functions of every individual in all generations are the same, i.e. $U_i^*\left(\cdot\right)=U^*\left(\cdot\right)$ for all ω_i. Next we show $U^*\left(\omega_0\right)$ satisfies the functional equation (10).

$$\begin{aligned}U^*\left(\omega_0\right)&\geqslant u\left(\omega_0\left(1-\lambda f_0\right)\right)+\beta f_0^{1-\epsilon}E_0\sum_{t=1}^{\infty}\beta^{t-1}\prod_{j=1}^{t-1}f_j^*\left(\omega^j\right)^{1-\epsilon}u\left(\omega_t\left(1-\lambda f_t^*\left(\omega^t\right)\right)\right)\\&=u\left(\omega_0\left(1-\lambda f_0\right)\right)+\beta f_0^{1-\epsilon}E_0\left[U^*\left(\omega_1\right)|\omega_0\right]\end{aligned}$$

for any given f_0 if $\left\{f_j^*\left(\omega^j\right)\right\}_{j=1}^{\infty}$ is the optimal solution that attains $U^*\left(\omega_1\right)$, the welfare of an individual living in generation 1 with ability ω_1. There exists a feasible plan that attains $U^*\left(\omega_0\right)$, so

$$U^*\left(\omega_0\right)=\max_{f_0\in[0,\frac{1}{\lambda}]}u\left(\omega_0\left(1-\lambda f_0\right)\right)+\beta f_0^{1-\epsilon}E_0\left[U^*\left(\omega_1\right)|\omega_0\right]$$

Hence $U^*\left(\cdot\right)$ satisfy the functional equation (10). Then we show that $U\left(\omega\right)$ is the maximum of the sequential problem for any given ω.

$$\begin{aligned}
U(\omega_0) &= \max_{f\in[0,\frac{1}{\lambda}]} u(\omega_0(1-\lambda f)) + \beta f^{1-\epsilon} E_0 [U(\omega_1)|\omega_0] \\
&\geqslant u(\omega_0(1-\lambda f_0(\omega_0))) + E_0 \begin{pmatrix} \beta f_0(\omega_0)^{1-\epsilon} u(\omega_1(1-\lambda f_1(\omega^1))) \\ +\beta^2 f_0(\omega_0)^{1-\epsilon} E_0 f_1(\omega^1)^{1-\epsilon} E[U(\omega_2)|\omega_1] \end{pmatrix} \\
&\geqslant \cdots \\
&\geqslant E_0 \sum_{t=0}^{T} \beta^t \prod_{j=0}^{t-1} f_j(\omega^j)^{1-\epsilon} u(\omega_t(1-\lambda f_t(\omega^t))) + \beta^{T+1} E_0 \prod_{j=0}^{T} f_j(\omega^j)^{1-\epsilon} U(\omega_{T+1})
\end{aligned}$$

for all feasible plan $\{f_t(\omega^t)\}_{t=0}^{\infty}$.

$$\beta^{T+1} \prod_{j=0}^{T} f_j(\omega^j)^{1-\epsilon} U(\omega_{T+1}) \leq \beta^{T+1} \left(\frac{1}{\lambda}\right)^{(1-\epsilon)(T+1)} \frac{u(\omega_{\max})}{1-\beta\lambda^{\varepsilon-1}}$$

$$\lim_{T\to\infty} \beta^{T+1} \left(\frac{1}{\lambda}\right)^{(1-\epsilon)(T+1)} = 0$$

Since there exists a feasible plan that attains $U(\omega_0)$,

$$U(\omega_0) = \max_{\{f_t(\omega^t)\}_{t=0}^{\infty}\in\left[0,\frac{1}{\lambda}\right]} E_0 \sum_{t=0}^{\infty} \beta^t \prod_{j=0}^{t-1} f_j(\omega^j)^{1-\epsilon} u(\omega_t(1-\lambda f_t(\omega^t)))$$

Therefore,

$$U(\omega_0) = U^*(\omega_0)$$

Proof of Proposition 5 (i) In this case, Equation (11) can be written as $\frac{f(\omega)^{\epsilon}}{(1-\lambda f(\omega))^{1-\sigma}} = A\omega^{-\sigma}$ where A is a constant. Using the implicit function theorem, it follows that $f' = -\frac{\sigma/\omega}{\frac{\epsilon}{f(\omega)}+\frac{\lambda(1-\sigma)}{1-\lambda f(\omega)}} < 0$.

Proof of Proposition 5 (ii) In deterministic case, $\omega_{t+1} = \omega_t$, $c_{t+1}(\omega_{t+1}) = c_t(\omega_t)$ for all

$\omega_t \in \Omega$ and equation (13) simplifies to:

$$f^{*\epsilon} = \beta \left(\frac{1}{\lambda} + \frac{1-\epsilon-\sigma}{\sigma} \left(\frac{1}{\lambda} - f^{*} \right) \right) \tag{19}$$

The left hand side of equation (19) is strictly increasing in f^* while the right hand side is strictly decreasing in f^*. Obviously $f^* > 0$. An interior solution with $f^* < 1/\lambda$ exists since $\lambda^{1-\epsilon} > \beta$.

Proof of Proposition 5 (iii) Let f^* denotes the optimal fertility given ω. Plug functional form of $u(\cdot)$ into equation (12)

$$U(\omega) = h(f^*)\,\omega^{\sigma} \tag{20}$$

where

$$h(f^*) \equiv \frac{1}{\sigma}(1-\lambda f^*)^{\sigma} + \frac{1}{1-\epsilon}\lambda f^* (1-\lambda f^*)^{\sigma-1} \tag{21}$$

We make a guess on the value function and let it take the form: $U(\omega) = A\omega^{\sigma}$ where A is a constant, independent of ω. Equating this guess with (20) results in:

$$A = h(f^*) \tag{22}$$

Thus, in order for A to be independent of ω, we must verify that the results f^* is independent of ω. Notice that,

$$E\left[U(\omega')\,|\omega\right] = E\left[A\omega'^{\sigma}|\omega\right] = A\omega^{\sigma} e^{\frac{\sigma^2\sigma_\epsilon^2}{2}}$$

The last equality holds because the assumption that ω' is lognormal distributed with $\ln\omega$ and σ_ε as the mean and variance of $\ln\omega'$. Plug this equality into (11) to obtain:

$$\lambda(1-\lambda f^*)^{\sigma-1}\omega^{\sigma} = A\beta(1-\epsilon)\, f^{*-\epsilon} e^{\frac{\sigma^2\sigma_\epsilon^2}{2}}\omega^{\sigma}$$

ω cancels out of this equation and therefore f^* is independent of ω confirming our guess. This expression together with (21) and (22) gives a rule to solve the optimal fertility f^*.

$$\frac{\lambda\sigma\left(1-\lambda f^*\right)^{\sigma-1}}{\beta\left(1-\epsilon\right)f^{*-\epsilon}}e^{-\frac{\sigma^2\sigma_\epsilon^2}{2}}=\left(1-\lambda f^*\right)^{\sigma}+\frac{f^*}{1-\epsilon}\lambda\sigma\left(1-\lambda f^*\right)^{\sigma-1}$$

To see that this equation has a solution, divide both sides by $\left(1-\lambda f^*\right)^{\sigma-1}$ to obtain

$$\frac{\lambda\sigma f^{*\epsilon}}{\beta\left(1-\epsilon\right)}\left[e^{-\frac{\sigma^2\sigma_\epsilon^2}{2}}-\beta f^{*1-\epsilon}\right]=1-\lambda f^*$$

Proof of Proposition 6 Part (i) directly applies Proposition 1 and Proposition 5 (i). To prove part (ii), we can apply Proposition 5 (ii), in which fertility is independent of ability when M is identity. We use this result to prove the distribution of every period as well as the limit distribution is the same with the initial one.

$$\begin{aligned}\pi_{t+1}\left(\omega_j\right)&=\frac{P_t}{P_{t+1}}\sum_{\omega_i\in\Omega}f\left(\omega_i\right)\pi_t\left(\omega_i\right)M(\omega_j,\omega_i)\\&=\frac{P_tf}{P_{t+1}}\sum_{\omega_i\in\Omega}\pi_t\left(\omega_i\right)M(\omega_j,\omega_i)\\&=\sum_{\omega_i\in\Omega}\pi_t\left(\omega_i\right)M(\omega_j,\omega_i)=\pi_t\left(\omega_j\right)\end{aligned}$$

The last equality holds because M is an identity matrix. So $\pi_t\left(\omega\right)=\pi_0\left(\omega\right)$ for all ω and all t. By taking limit we have $\pi^*\left(\omega\right)=\pi_0\left(\omega\right)$. Part (iii) follows Proposition 5 (iii). The conditional variance of $\ln\omega_t$ diverges to infinite because $\ln\omega_t=\ln\omega_0+\sum_{i=0}^{t-1}\epsilon_i$, $E\left(\ln\omega_t|\omega_0\right)=\ln\omega_0$, $Var\left(\ln\omega_t|\omega_0\right)=t^2\sigma_\epsilon^2$ and $\lim_{t\to\infty}Var\left(\ln\omega_t|\omega_0\right)=\infty$.

Proof of Proposition 7. Notice that

$$\begin{aligned}
U(\omega) &= \max_{f_t \in [0,1/\lambda]} u((1-\lambda f)\omega) + \beta f^{1-\epsilon} E\left[U(\omega')|\omega\right] \\
&\geq \max_{[\underline{f}(\omega),\overline{f}(\omega)]} u((1-\lambda f)\omega) + \beta f^{1-\epsilon} E\left[U(\omega')|\omega\right] := U^1(\omega') \\
&\geq \max_{[\underline{f}(\omega),\overline{f}(\omega)]} u((1-\lambda f)\omega) + \beta f^{1-\epsilon} E\left[U^1(\omega')|\omega\right] := U^2(\omega') \\
&.. \\
&\geq \max_{[\underline{f}(\omega),\overline{f}(\omega)]} u((1-\lambda f)\omega) + \beta f^{1-\epsilon} E\left[U^r(\omega')|\omega\right] = U^r(\omega')
\end{aligned}$$

where the first inequality is strict if a constraint is binding for any particular ω, the remaining inequalities follow from the contraction mapping recursion, and the final inequality uses the contraction mapping theorem. Furthermore, a strict inequality for a particular ω translates into a strict inequality for all $\omega' s$ since M is a regular Markov chain meaning that, regardless of initial ability there is positive probability that someone in the dynasty will reach a binding state in finite time. The second part of the proposition follows because fertility restrictions do not change the marginal costs of having children but it decreases the marginal benefits by reducing $U(\omega)$ for all ω (see equation (11)). Hence an upper bound of fertility makes people have fewer children than (or the same number of children with) the unrestricted case.

Proof of Proposition 8 By (1) and Proposition 7,

$$\begin{aligned}
P_1(\omega_j) &= \sum_{\omega_i \in \Omega} f(\omega_i) P_0(\omega_i) M(\omega_j,\omega_i) \\
&\geq \sum_{\omega_i \in \Omega} f^r(\omega_i) P_0^r(\omega_i) M(\omega_j,\omega_i) = P_1^r(\omega_j)
\end{aligned}$$

where $P_0(\omega_i) = P_0^r(\omega_i)$. An inductive argument guarantees

$$P_{t+1}(\omega_j) = \sum_{\omega_i \in \Omega} f(\omega_i) P_t(\omega_i) M(\omega_j, \omega_i) \geq \sum_{\omega_i \in \Omega} f^r(\omega_i) P_t^r(\omega_i) M(\omega_j, \omega_i) = P_{t+1}^r(\omega_j)$$

for all ω_j and all finite $t \geq 0$.

$$\begin{aligned} W^r(\beta_p) &= \sum_{t=0}^{\infty} \sum_{\omega \in \Omega} \beta_p(t) U^r(\omega) P_t^r(\omega) \\ &\leq \sum_{t=0}^{\infty} \sum_{\omega \in \Omega} \beta_p(t) U(\omega) P_t^r(\omega) \\ &\leq \sum_{t=0}^{\infty} \sum_{\omega \in \Omega} \beta_p(t) U(\omega) P_t(\omega) = W(\beta_p) \end{aligned}$$

Proof of Proposition 9 By the proof of Proposition 1(ii),

$$\pi_{t+1}(\omega_j) = \sum_{\omega_i \in \Omega} \pi_t(\omega_i) M(\omega_j, \omega_i)$$

If $\pi_0(\omega_i) = \mu(\omega_i)$, then

$$\pi_1(\omega_j) = \sum_{\omega_i \in \Omega} \mu(\omega_i) M(\omega_j, \omega_i) = \mu(\omega_j),$$

and by iteration, $\pi_t(\omega) = \mu(\omega)$for all t and all ω. So restriction on fertility upper bound only reduces individual utility by Proposition 8 but does not affect the ability distribution. It decreases social welfare defined by (3).

Proof of Proposition 10 This Proposition relies on Proposition 7 (ii)'s results in $\pi_t(\omega) = \pi^*(\omega) = \pi_0(\omega)$ when M is an identity. Similar with Proposition 10, restriction does not alter distribution, which together with Proposition 8, finishes the proof.

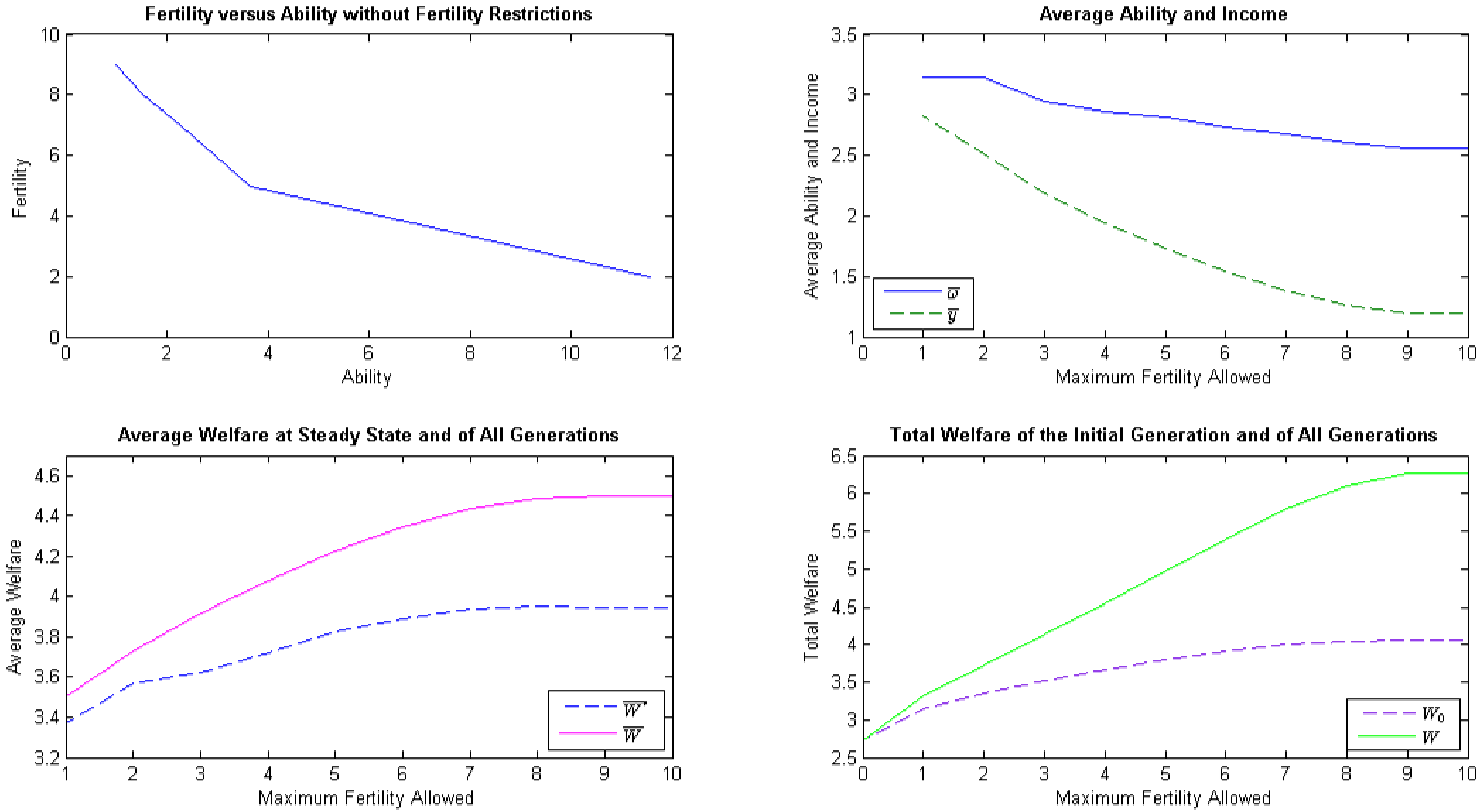

Figure 1. Effects of Fertility Restrictions: Upper Bound

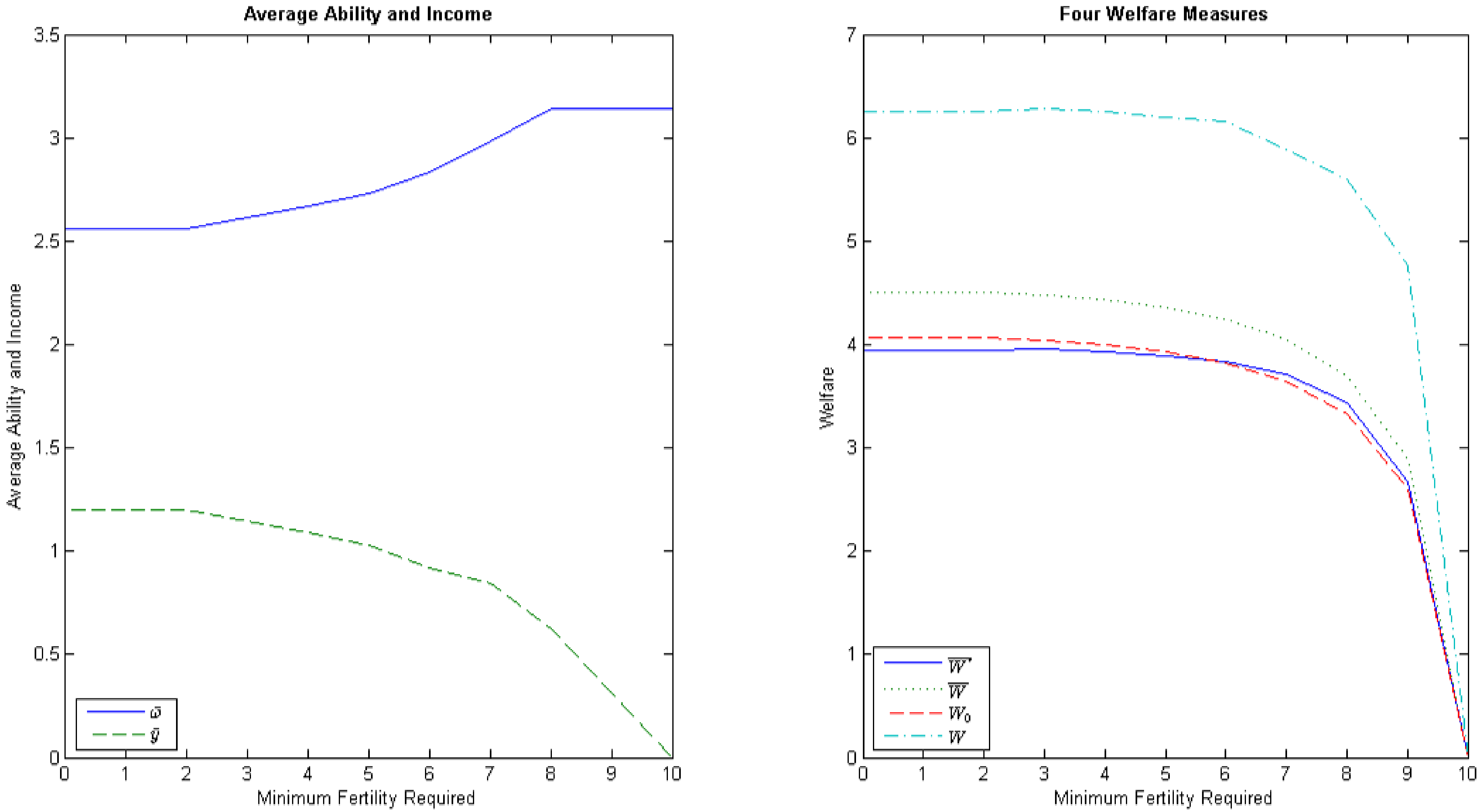

Figure 2. Effects of Fertility Restrictions: Lower Bound

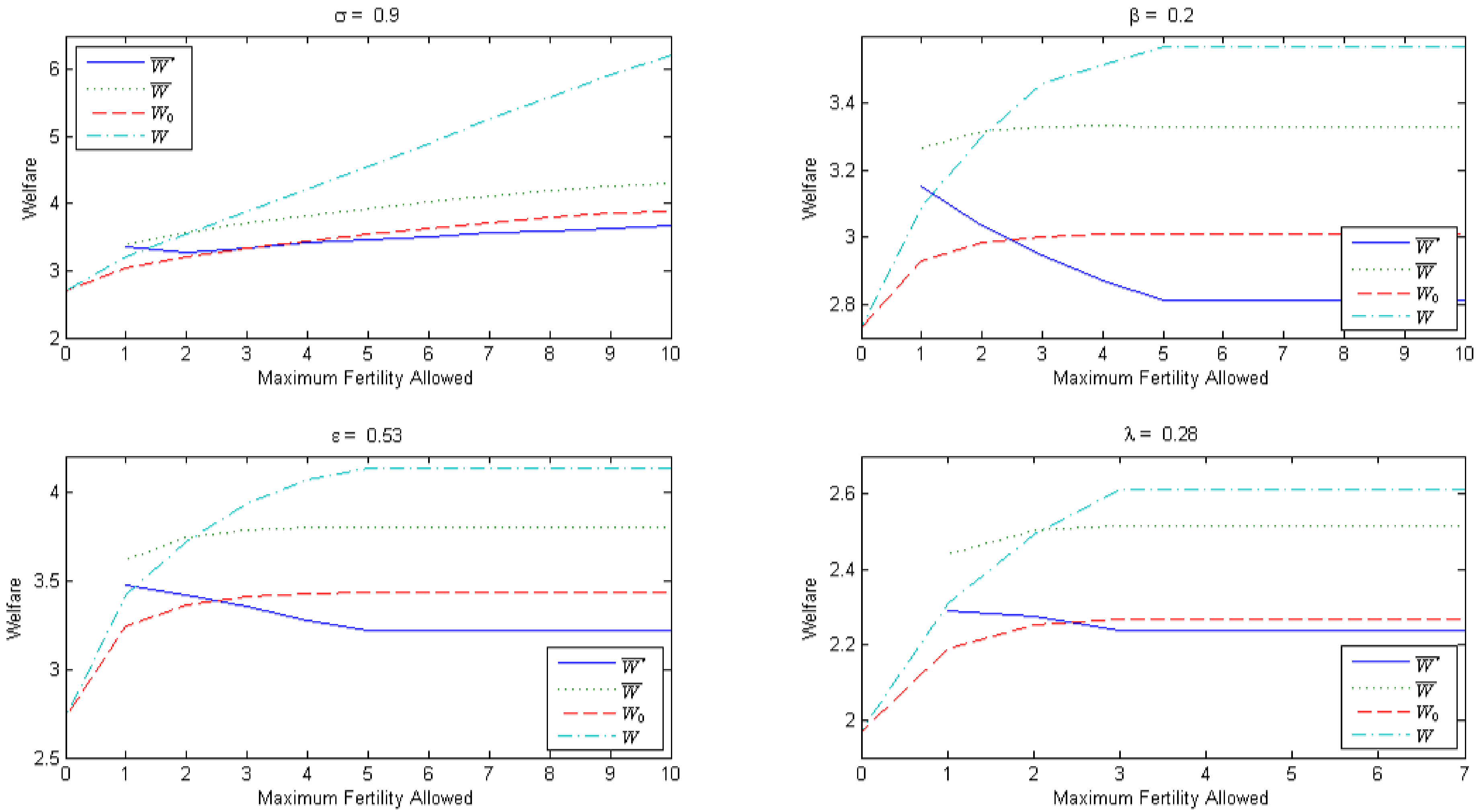

Figure 3. Robustness Checks