

Discussion of “The Dynamics of Gender Norms and Fertility Decline”

Xincheng Qiu
Peking University

July 20, 2026

[Xi and Zhou \(2026\)](#) study why fertility has declined so rapidly in many economies. They develop a quantitative overlapping-generations model in which fertility is shaped by two interacting forces. First, gender-biased technological change raises women’s relative wages and therefore the opportunity cost of childbearing. This lowers desired fertility. Second, social norms governing the gender division of childcare adjust sluggishly, preventing households from optimally reallocating childcare time from women to men. This further raises the shadow price of children and amplifies the decline in fertility. The model is calibrated to South Korea and is used to conduct several counterfactual exercises. A particularly striking result is that subsidizing husbands’ childcare yields larger long-run fertility gains than subsidizing wives’ childcare because it accelerates the transition toward a more egalitarian allocation of childcare.

This is an inspiring paper that tackles an important question. The mechanism is clear, and the authors makes a thoughtful attempt to quantify it despite the measurement challenges. My discussion develops two themes. First, I discuss what I see as the paper’s most portable contribution: its framework for social-norm formation. I isolate the mechanism in a stripped-down environment, characterize its dynamics, and highlight its key insights. Second, I discuss directions for future research, including applying the norm-formation framework to other settings, identifying the behavior to which the norm attaches, and refining several aspects of the model.

1 Social Norm Formation

How norms form, persist, and change is among the oldest questions in the social sciences. Economists have long recognized that behavior is shaped not only by prices and incomes, but also by the expectations, approval, and sanctions of others. The economics literature

offers at least four broad mechanisms. The first is coordination and evolutionary selection, in which norms emerge as equilibria of recurrent interactions and evolve through stochastic experimentation and adaptive best responses (Young, 1993; Burke and Young, 2011; Young, 2015). The second is social image and signaling, in which individuals conform because their actions affect how others perceive them, so public conformity may persist despite heterogeneous private preferences (Bernheim, 1994; Bursztyn and Jensen, 2017). The third is intergenerational cultural transmission, pioneered by Bisin and Verdier (2000, 2001, 2011, 2023), in which parents exert costly socialization effort to shape their children’s traits under imperfect empathy. The fourth is social learning and information diffusion, in which beliefs about the consequences of a behavior are updated from the experiences of others, and behavior and attitudes change as information accumulates (Fogli and Veldkamp, 2011; Fernández, 2013).

Xi and Zhou (2026) are closest to the third tradition but depart from it in a way that facilitates quantitative work. Bisin and Verdier (2000, 2001) place a forward-looking socialization decision at the center, with parents anticipating how their effort shapes the next generation. This paper instead gives older cohorts a per-period opinion-expression problem. This preserves an intergenerational link between past practices and current behavior while keeping the norm dynamics tractable. The prevailing norm is a population-weighted average of the opinions expressed by living older cohorts. Each older cohort chooses its expressed opinion by balancing the contemporary economic environment against the practices it adopted when young. The current cohort, in turn, finds it costly to deviate from the prevailing norm. Two parameters govern these forces: a stubbornness parameter governing how strongly older cohorts’ expressed opinions remain anchored to their own past behavior, and a social-pressure parameter governing how costly it is for the current cohort to deviate from the prevailing norm. Together, these forces make norms a source of intergenerational inertia.

1.1 General Setup

Here I present a stripped-down environment. Time is discrete and indexed by t . Let x_t denote the practice to which the norm attaches. The economic fundamental determines the frictionless ideal x_t^* , defined as the practice the active cohort would choose in the absence of social pressure. In Xi and Zhou (2026), x_t is the ratio of mothers’ to fathers’ childcare time, and changes in women’s relative wages determine the ideal x_t^* . The prevailing norm η_t is generated by two behavioral blocks and one aggregation rule.

The active cohort. The cohort currently choosing its behavior trades off the private cost of departing from its ideal against the disutility of deviating from the prevailing norm:

$$x_t = \arg \min_x \frac{1}{2} (x - x_t^*)^2 + \frac{\lambda}{2} (x - \eta_t)^2.$$

Quadratic losses make the mechanics transparent, but the same formulation arises as a local approximation under general smooth, convex losses. The first-order condition gives

$$x_t = (1 - \theta) x_t^* + \theta \eta_t, \quad \theta \equiv \frac{\lambda}{1 + \lambda}. \quad (1)$$

Here $\theta \in [0, 1]$ is the pass-through of the prevailing norm into current behavior. When $\lambda = 0$ ($\theta = 0$), the active cohort ignores the norm and chooses its frictionless ideal; as $\lambda \rightarrow \infty$ ($\theta \rightarrow 1$), the active cohort conforms fully to the norm.

Older cohorts. An older cohort that chose x_{t-j} when young forms the opinion it expresses at date t by balancing the current ideal against its own past practice:

$$\tilde{\eta}_{j,t} = \arg \min_{\eta} \frac{1}{2} (\eta - x_t^*)^2 + \frac{\psi}{2} (\eta - x_{t-j})^2,$$

where the vintage distance j denotes the number of periods since the cohort made its active-period choice. The first-order condition gives

$$\tilde{\eta}_{j,t} = (1 - s) x_t^* + s x_{t-j}, \quad s \equiv \frac{\psi}{1 + \psi}. \quad (2)$$

Here $s \in [0, 1]$ is the pass-through of a cohort's own history into its expressed opinion. When $\psi = 0$ ($s = 0$), older cohorts fully re-evaluate their views in light of current conditions; as $\psi \rightarrow \infty$ ($s \rightarrow 1$), they express only the practices they adopted when young.

The prevailing norm. The prevailing norm is the weighted average of expressed opinions across the set $\mathcal{J}_t$ of living older cohorts:

$$\eta_t = \sum_{j \in \mathcal{J}_t} \omega_{j,t} \tilde{\eta}_{j,t}, \quad (3)$$

where $\omega_{j,t} \geq 0$ is the weight assigned to vintage j , capturing its influence on the aggregate norm, and $\sum_j \omega_{j,t} = 1$. In [Xi and Zhou \(2026\)](#), these weights are given by cohort population shares and are therefore endogenous to the history of fertility.

1.2 Norm Dynamics

Substituting the opinion rule (2) into the aggregation rule (3) and using $\sum_j \omega_{j,t} = 1$ gives

$$\eta_t = (1 - s)x_t^* + s\bar{x}_t, \quad (4)$$

where $\bar{x}_t \equiv \sum_j \omega_{j,t} x_{t-j}$ is the weighted history of past practices. The first term captures contemporary re-evaluation, and the second term captures norm persistence arising from past behavior, making the paper's two channels of norm dynamics explicit. Replacing each x_{t-j} by its reaction (1) expresses the norm as a distributed lag of past ideals and past norms:

$$\eta_t = (1 - s)x_t^* + s \sum_{j \in \mathcal{J}_t} \omega_{j,t} \left[(1 - \theta)x_{t-j}^* + \theta \eta_{t-j} \right].$$

This representation shows that the prevailing norm inherits persistence from past norms through the feedback of social pressure into earlier cohorts' behavior.

Substituting (4) into the behavior rule (1) gives the law of motion for behavior:

$$x_t = (1 - \phi)x_t^* + \phi\bar{x}_t, \quad (5)$$

where

$$\phi \equiv s \cdot \theta = \frac{\lambda}{1 + \lambda} \frac{\psi}{1 + \psi}.$$

Behavior is a convex combination of the current ideal and a weighted average of past behavior, with the weight ϕ on history equal to the product of the two pass-throughs. Inertia requires both frictions. If $\lambda = 0$, the young ignore the norm; if $\psi = 0$, the old fully re-evaluate. In either case, $\phi = 0$ and behavior tracks the ideal without lag.

An intergenerational adjustment-cost interpretation. The economy generates a behavior path that is observationally equivalent to that of a representative agent who solves, in each period,

$$x_t = \arg \min_x \frac{1}{2}(x - x_t^*)^2 + \frac{\kappa}{2}(x - \bar{x}_t)^2, \quad \kappa = \frac{\phi}{1 - \phi} = \frac{\lambda\psi}{1 + \lambda + \psi}.$$

To see this, the first-order condition is $(x - x_t^*) + \kappa(x - \bar{x}_t) = 0$, which implies $x_t = \frac{1}{1 + \kappa}x_t^* + \frac{\kappa}{1 + \kappa}\bar{x}_t$. Setting $\kappa = \phi/(1 - \phi)$ recovers equation (5). The equivalent adjustment-cost parameter κ is symmetric in λ and ψ and vanishes when either parameter is zero.

1.3 Transitional Wedge

Define the wedge between the frictionless ideal and realized behavior as

$$\Delta_t \equiv x_t^* - x_t.$$

Subtracting (1) from x_t^* gives $\Delta_t = \theta(x_t^* - \eta_t)$. Subtracting (4) from x_t^* gives $x_t^* - \eta_t = s(x_t^* - \bar{x}_t)$. Since the weights sum to one,

$$x_t^* - \bar{x}_t = \sum_j \omega_{j,t}(x_t^* - x_{t-j}) = \sum_j \omega_{j,t}[(x_t^* - x_{t-j}^*) + \Delta_{t-j}].$$

Combining these expressions yields

$$\Delta_t = \phi \sum_j \omega_{j,t}[(x_t^* - x_{t-j}^*) + \Delta_{t-j}]. \quad (6)$$

The wedge is a weighted average of past wedges, additionally driven by the distance the frontier has traveled over each cohort's vintage.

We now consider a frontier that advances at a constant rate, $x_t^* = x_0^* + \mu t$, so that $x_t^* - x_{t-j}^* = \mu j$. Consider a balanced growth path (BGP) along which the age structure is stationary and the wedge is constant. Along the BGP, equation (6) reduces to

$$\Delta = \phi \sum_j \omega_j [\mu j + \Delta],$$

which implies

$$\Delta = \frac{\phi}{1 - \phi} \mu \bar{j} = \frac{\lambda \psi}{1 + \lambda + \psi} \mu \bar{j}, \quad (7)$$

where $\bar{j} \equiv \sum_j \omega_j j$ denotes the mean age, or mean vintage distance, of the cohorts whose opinions determine the norm. Thus, the economy permanently lags behind a moving frontier. The absolute value of the wedge is larger when the frontier moves faster ($|\mu| \uparrow$), opinion holders are further behind the frontier ($\bar{j} \uparrow$), or social pressure and stubbornness are stronger ($\lambda, \psi \uparrow$).

Equation (7) captures the key insight of the application in [Xi and Zhou \(2026\)](#). Rising female relative wages imply a declining frictionless childcare ratio ($\mu < 0$). The resulting negative wedge ($\Delta < 0$) indicates that households consistently allocate more childcare time to mothers than is privately optimal ($x_t > x_t^*$). This forces families to use the mother's increasingly expensive time, raising the shadow price of children and reducing fertility below its frictionless level. Stronger social pressure (λ) and stubbornness (ψ) enlarge this wedge and therefore further depresses fertility. Moreover, as fertility declines, older cohorts account for a larger share of

the population, increasing the mean vintage distance $\bar{j}$ and thereby widening the wedge. This endogenous demographic response further reduces fertility, creating a vicious cycle between population aging and sluggish norm adjustment.

2 Future Directions

2.1 Potential Applications

The mechanism is not specific to fertility. It applies whenever an economic fundamental shifts the privately optimal practice, current decision-makers face a cost of departing from a socially enforced standard, and influential older cohorts remain partly anchored to their own past conduct. An application can therefore be organized around two parameters (social pressure λ and stubbornness ψ) and three objects (the practice x_t , the frictionless ideal x_t^* , and the norm η_t). Here I list two examples.

Scientific paradigms. Consider the adoption of a new method or paradigm within a scientific discipline. This is the setting of [Planck's](#) remark that science advances one funeral at a time (see, e.g., [Azoulay et al., 2019](#), for empirical evidence). The practice x_t is the extent to which researchers use the new method. The ideal x_t^* is the degree of adoption warranted by the current frontier in data, computation, and theory. The norm η_t is the discipline's prevailing standard of what constitutes legitimate work, enforced through peer review, grant allocation, promotion, and hiring. Social pressure λ then has a direct career interpretation, since researchers who depart from prevailing standards risk rejection at journals, denial of grants and jobs, and professional marginalization. The stubbornness parameter ψ measures the extent to which senior scholars evaluate new work through the paradigm in which they were trained.

Pension reform. Consider a government choosing the statutory retirement age in response to rising life expectancy and population aging. The practice x_t is the retirement age actually implemented, and the frictionless ideal x_t^* is the retirement age that would balance pension-system sustainability, labor supply, and workers' health under current demographic and economic conditions. The norm η_t is society's prevailing view of the age at which people ought to stop working and become entitled to retirement benefits. Social pressure λ arises because pension reform must be seen as legitimate by workers, unions, and employers. A government that raises the retirement age far above the prevailing norm may face electoral resistance, protests, or demands for compensating transfers (see, e.g., [Boeri et al., 2002](#), for empirical evidence). Stubbornness ψ arises because cohorts remain anchored to the system under which they entered the

labor market and planned their careers, and may regard the retirement age promised during their working lives as an entitlement.

2.2 The Content of the Norm

The paper models the relevant norm as governing the gender division of childcare, specifically the ratio of mothers' to fathers' childcare time. This choice is not innocuous. Because the norm constrains the reallocation of childcare time, sluggish norm adjustment amplifies the fertility decline in response to rising female wages. However, if the norm instead governed fertility itself, for example, through a social expectation about the appropriate number of children, then sluggish norm adjustment would dampen the decline, since a traditional norm favoring larger families would resist downward pressure on fertility. Thus, the same machinery of norm formation can produce opposite implications depending on the behavior to which the norm attaches. In this paper, the choice is supported by [Myong et al. \(2021\)](#), whose structural estimates suggest that childcare norms appear to matter more than marriage norms in South Korea. More generally, identifying the content of the norm is an important direction for future research.

How might the content of the norm be identified? This is a difficult question, but three approaches seem promising. First, survey instruments can directly distinguish among normative beliefs, such as beliefs about gender roles and ideal family size. Second, variations across policies may be informative about the norms they target. For example, pro-natalist campaigns may operate directly on the family-size norm, while paternity-leave mandates may operate primarily on the childcare-division norm. Third, a structural model could nest both norms and discipline them jointly with data on fertility, time use in childcare, and attitudes. Since this problem recurs across applications, a useful research agenda would develop systematic criteria for selecting among candidate normative objects.

2.3 Opinion Aggregation and Other Mechanisms

The simplicity of the framework is a virtue for the present quantitative application. A natural extension within the current structure concerns how expressed opinions are aggregated into the prevailing norm.

The model uses a population-weighted average of cohort opinions. This raises two questions. First, what should determine the relevant weights? Should they reflect population, as in the baseline, or political power, media visibility, or institutional authority? The scientific-paradigms application above suggests that influence may depend on who controls access to journals, fund-

ing, and employment. Second, what should the aggregation protocol be? Do norms reflect the average or median opinion, the loudest voice, or the views of a small group with disproportionate authority? Different protocols generate different norm dynamics. A median-based norm is robust to outliers, whereas a norm shaped by a vocal minority may shift abruptly. Alternative aggregation rules could therefore help explain why norms sometimes change suddenly rather than gradually, as in social movements.

Other mechanisms would require more substantial departures. Opinion expression could be strategic. For example, a grandparent who wants grandchildren might voice more progressive childcare views than they privately hold, if doing so lowers the effective cost of children for their adult offspring. Cohorts could also learn from one another’s opinions and experiences. These mechanisms are well established elsewhere in the literature, but incorporating them here would require modeling strategic communication or belief updating and is therefore better viewed as a broader direction for future work.

References

- AZOULAY, P., C. FONS-ROSEN, AND J. S. G. ZIVIN (2019): “Does science advance one funeral at a time?” *American Economic Review*, 109, 2889–2920.
- BERNHEIM, B. D. (1994): “A theory of conformity,” *Journal of Political Economy*, 102, 841–877.
- BISIN, A. AND T. VERDIER (2000): ““Beyond the melting pot”: cultural transmission, marriage, and the evolution of ethnic and religious traits,” *The Quarterly Journal of Economics*, 115, 955–988.
- (2001): “The economics of cultural transmission and the dynamics of preferences,” *Journal of Economic Theory*, 97, 298–319.
- (2011): “The economics of cultural transmission and socialization,” in *Handbook of Social Economics*, Elsevier, vol. 1, 339–416.
- (2023): “Advances in the economic theory of cultural transmission,” *Annual Review of Economics*, 15, 63–89.
- BOERI, T., A. BÖRSCH-SUPAN, AND G. TABELLINI (2002): “Pension reforms and the opinions of European citizens,” *American Economic Review*, 92, 396–401.
- BURKE, M. A. AND H. P. YOUNG (2011): “Social norms,” in *Handbook of Social Economics*, Elsevier, vol. 1, 311–338.

- BURSZTYN, L. AND R. JENSEN (2017): “Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure,” *Annual Review of Economics*, 9, 131–153.
- FERNÁNDEZ, R. (2013): “Cultural change as learning: The evolution of female labor force participation over a century,” *American Economic Review*, 103, 472–500.
- FOGLI, A. AND L. VELDKAMP (2011): “Nature or nurture? Learning and the geography of female labor force participation,” *Econometrica*, 79, 1103–1138.
- MYONG, S., J. PARK, AND J. YI (2021): “Social norms and fertility,” *Journal of the European Economic Association*, 19, 2429–2466.
- PLANCK, M. (1949): *Scientific autobiography and other papers*, New York: Philosophical Library.
- XI, X. AND A. ZHOU (2026): “The dynamics of gender norms and fertility decline,” *International Economic Review*.
- YOUNG, H. P. (1993): “The evolution of conventions,” *Econometrica: Journal of the Econometric Society*, 57–84.
- (2015): “The evolution of social norms,” *Annual Review of Economics*, 7, 359–387.