Technical Appendix 5: Generalizing the model

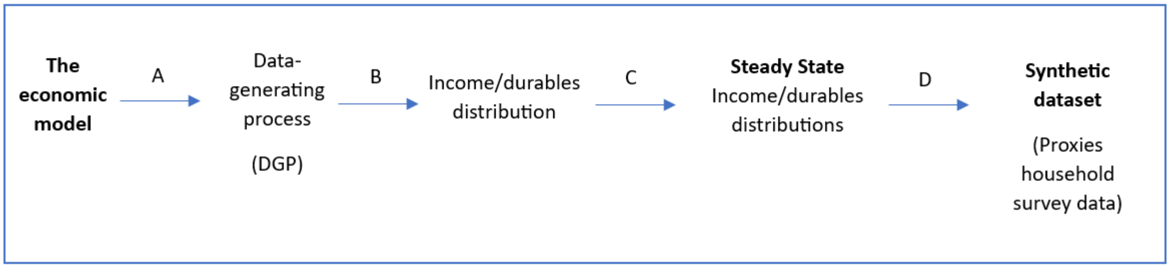


Fig. A

The infographic above (Fig. A) outlines the composite "generative" approach used in the paper to understand and interpret cluster formation in total durable ownership data. Below, we discuss ways to generalize the primary economic model, and trace how if at all the generalization might impact the interpretation of clusters in the synthetic dataset.

5.1 Understanding how the underlying data generating process (DGP) affects the income distribution (Step B in Fig. A)

Note that Section 2 presents a "generative" model, which can, in principle, be made very complex, e.g. by including a continuum of possible education expenditure decisions over a sequence of years and very complex joint distributions of market wage and marriage market outcomes. Such models would be computationally intense and take an extremely long time to solve. In the manuscript, therefore, we have chosen an approach that is arguably less general in terms of parameters (e.g. we have 2 wage levels in the labour and marriage markets, 2 education levels etc). The exposition below is intended to offer insights on which features of the resulting steady state distribution may indeed be influenced (or not) by the model simplifications. We hope this section will also offer direction while making potential generalizations of the model (presented in later sections).

A single channel of income generation

Consider 2 economies with the following DGPs for income in every time period t:

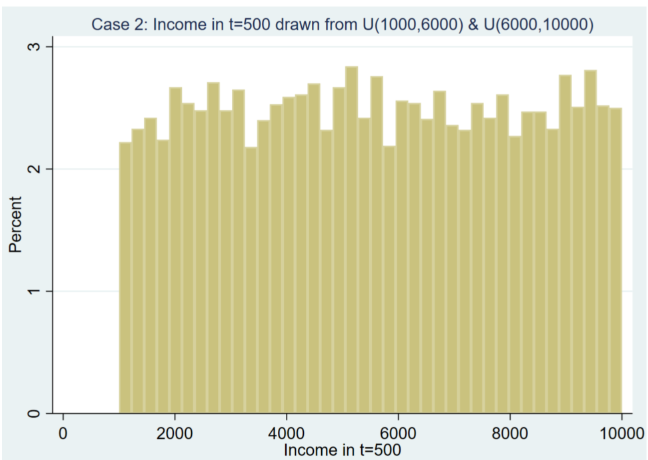
- Case 1 (DGP1): Economy 1, where total incomes I_t in any period t, are drawn from a uniform distribution over the range (a, b) [instead of incomes taking only 3 levels – low, medium or high – as in our model]
- Case 2 (DGP2): Economy 2, where all households with $I_t > d$ draw income $I_{t+1} \sim U(d, b)$ and all households with $I_t < d$ draw income $I_{t+1} \sim U(a,d)$, for some c

such that $a < d < b$. Thus, DGP2 embodies a "wealth-begets-wealth" mechanism whereby households that have income above d in any period are assured an income above d in the next period (and those with income below d end up with income below d in the next period).

Suppose, $a = 1000, d = 6000, b = 10000$.

Let us draw income data from DGP1 and DGP2, starting at time $t = 0$ and up to $t = 1000$. Consider the distribution of I_t in, say, period $t = 500$, in Economy 1 and Economy 2.

Fig. 1



Notice that the distribution of income I_t in $t = 500$ looks identical for Economy 1 and Economy 2, even though the underlying data generating processes are quite different in the 2 societies.

Now look at the the $two - period\ sum$ of incomes in $t = 499$ and $t = 500$ in Economy 1 and Economy 2:

Fig. 2

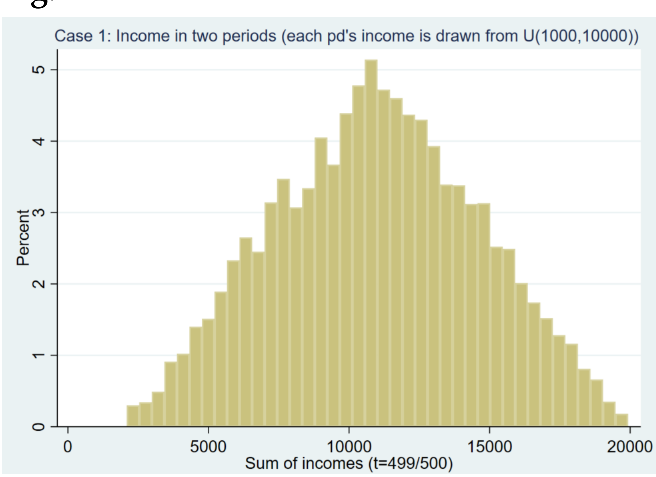
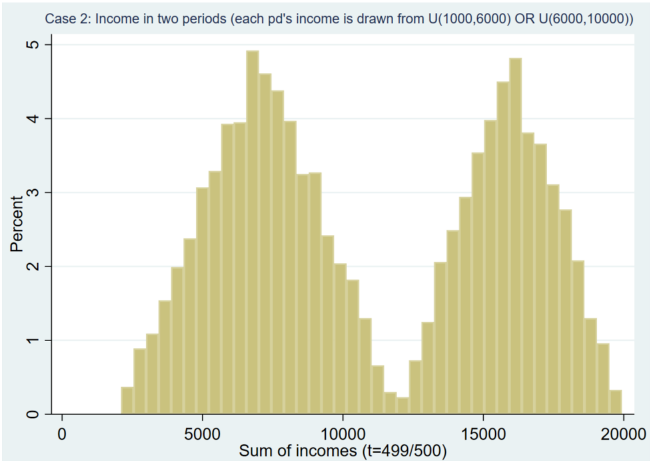


Fig. 3



Notice that even though there is no observable difference in the distribution of single-period incomes under DGP1 and DGP2, the sum of $two-period$ incomes looks very different for the 2 cases, viz. one cluster for Economy 1 and 2 clusters for Economy 2.

The illustration above demonstrates the clustering generated in $two-period$ incomes when a wealth-begets-wealth mechanism is embedded in DGP2.

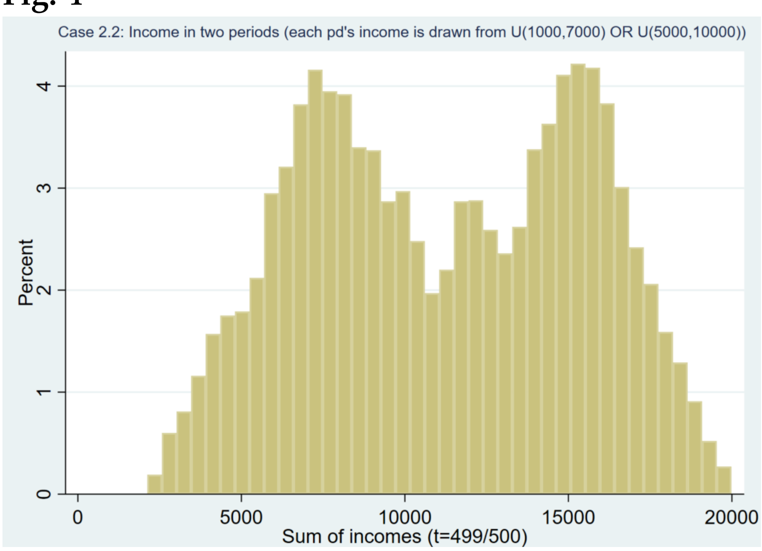
Now consider a variation of DGP2 (Case 2.2, DGP3) as follows:

- Case 2.2 (DGP2.2), Economy 3, where all households with $I_t > d$ draw income $I_{t+1} \sim U(d-\epsilon, b)$ and all households with $I_t < d$ draw income $I_{t+1} \sim U(a,d+\zeta)$, for some c such that $a < d < b$.

Suppose $a = 1000, d = 6000, b = 10000, \epsilon = \zeta = 1000$.

Let us now draw income data from DGP2.2, starting at time $t = 0$ and up to $t = 1000$. Consider the distribution of $(I_t + I_{t-1})$ in, say, period $t = 500$, in Economy 3.

Fig. 4

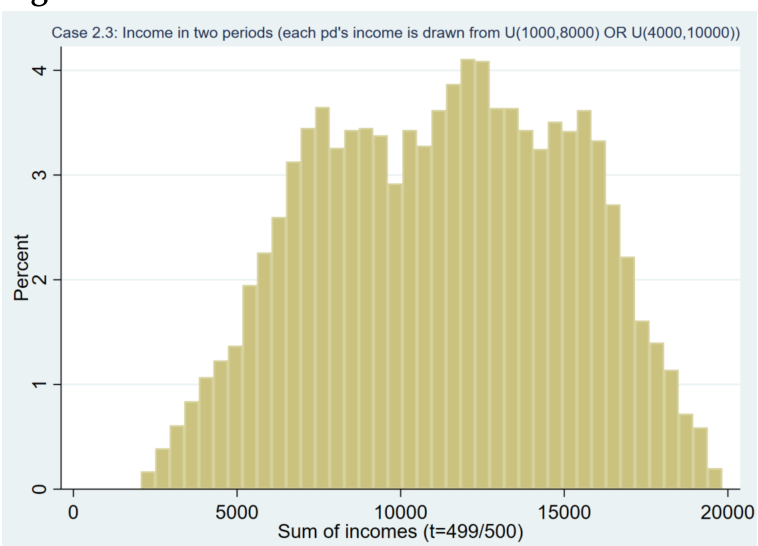


In the figure above, we see a third cluster emerging (with peak around 11000) due to the overlap of the distributions via ϵ and ζ. Recall that these clusters emerge in two-period incomes, even though incomes in each period are drawn from a uniform distribution. Now consider yet another DGP (DGP2.3, Economy 4) that allows greater overlap in the wealth-begets-wealth mechanism, i.e. $\epsilon = \zeta = 2000$.

- Case 2.3 (DGP2.3), Economy 4, where all households with $I_t > d$ draw income $I_{t+1} \sim U(d - \epsilon, b)$ and all households with $I_t < d$ draw income $I_{t+1} \sim U(a, d + \zeta)$, for some c such that $a < d < b$; $a = 1000, d = 6000, b = 10000$, $\epsilon = \zeta = 2000$.

Consider the distribution of $(I_t + I_{t-1})$ in, say, period $t = 500$, in Economy 3.

Fig. 5



Conclusions:

1. A wealth-begets-wealth mechanism generates clusters in two-period incomes (Figs. 3-5). We see clusters even when incomes are drawn from a *continuous* range of incomes (hence, the simplifying assumption of 3 income levels – L, M, H – in our model is not the driving force behind cluster generation).

2. The overlap in the wealth-begets-wealth distributions (i.e. values of ϵ, ζ) influence the number of clusters.

3. As overlaps (ϵ, ζ) become larger, the clusters "mix" (i.e. they are less discernible *visibly* even though the clustering mechanism is still at play). (In the limiting case where $\epsilon = (d - a), \zeta = (b - d)$, we return to DGP1 where there is no wealth-begets-wealth mechanism.)

Two channels of income generation

Suppose economies have 2 channels of income generation, and that the total income in any period is the sum of income generated across the 2 channels. Suppose that income draws in each channel are *independent*. Consider the following DGPs:

- Case 3 (DGP3): Economy 5, where, in each channel, incomes are drawn (independently) from $U(500, 5000)$ in every period (no wealth begets wealth in any channel).

- Case 4 (DGP4): Economy 6, where, in each channel, income is drawn (independently) from $U(3000, 5000)$ if income in the previous period is above 3000; income is drawn from $U(500, 3000)$ if income in the previous period is below 3000 (wealth begets wealth in *both* channels, independently).

As before, the distirbution of single period incomes drawn from DGP3 and DGP4 are indistinguishable:

Fig. 6

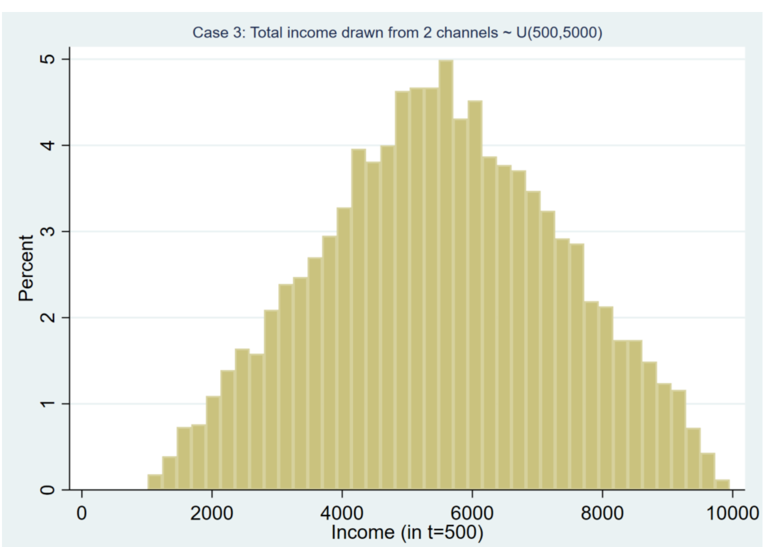
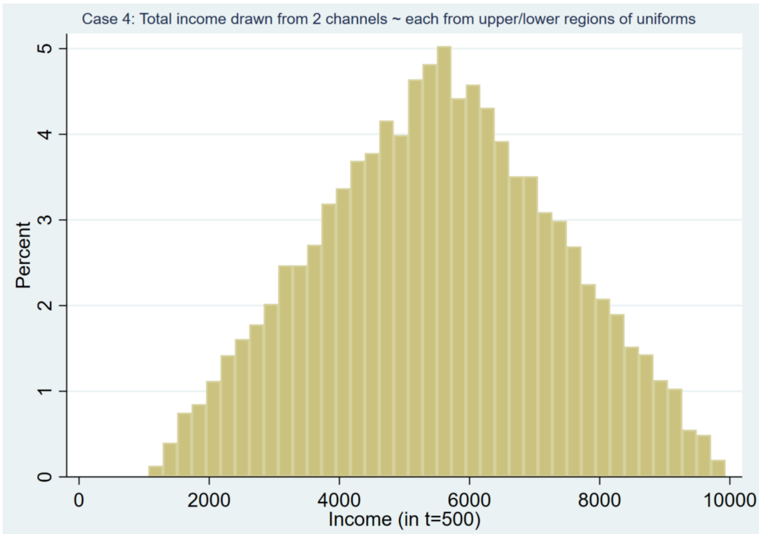


Fig. 7



However, the sum of two-period incomes looks very different under DGP3 and DGP4:

Fig. 8

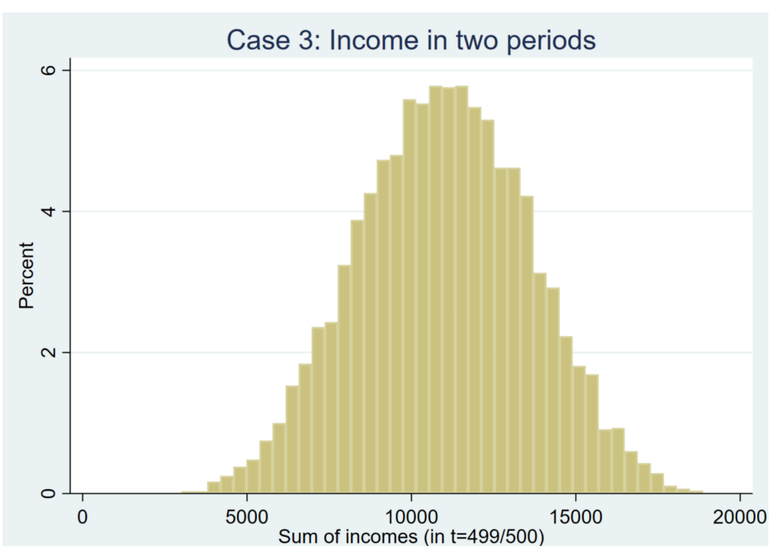
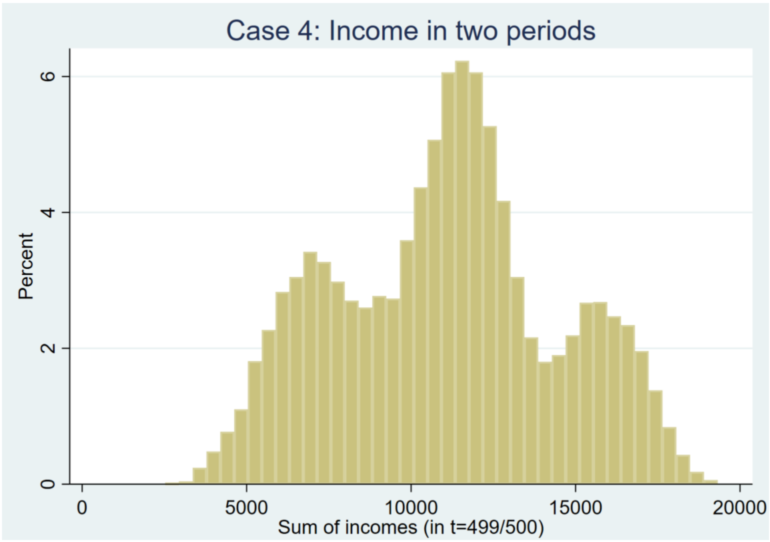


Fig. 9

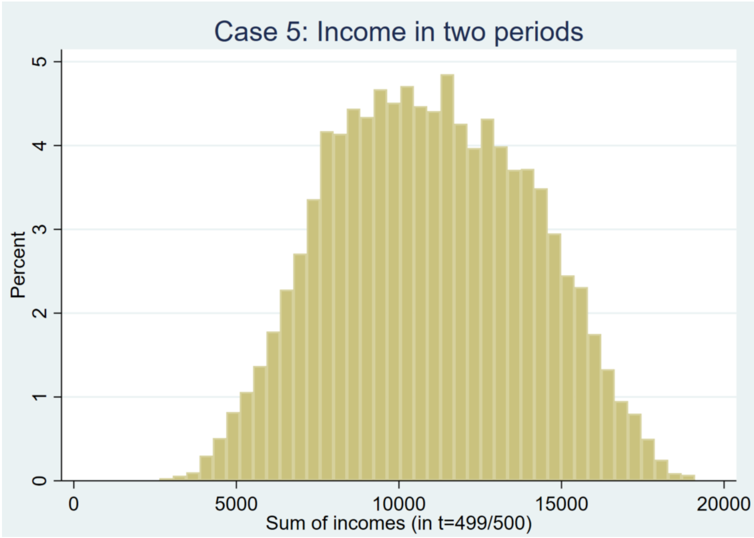


What if there is a wealth-begets-wealth mechanism in only 1 of the channels, as follows?

- Case 5 (DGP5): Economy 7, where, in the first channel, income is drawn from $U(500, 5000)$ in every period; in the second channel, income is drawn from $U(500, 3000)$ if income in the previous period (in this channel) is below 3000 (wealth begets wealth in *one* channel only, independent draws across channels).

Notice, in the distribution of two-period incomes drawn from DGP5, that the clustering mechanism in 1 channel of income generation is diluted by the (independent) effect of the other channel with no wealth-begets-wealth.

Fig. 10



Observe the differences in the shape of the distributions in Figs. 3-5 (one channel with varying "characteristics" of the wealth-begets-wealth mechanism), Fig. 9 (2 channels, independent, each with wealth-begets-wealth), Fig. 10 (2 channels, independent, only one channel has wealth-begets-wealth).

The above exposition demonstrates the process of cluster formation purely from generative DGPs – there is no economics or concept of steady state equilibrium in the above.

Let us now consider how the economic forces in our model (Section 2) might impact the DGP – *in steady state equilibrium* – and how these forces might influence cluster formation (such as in Figs. 1-10). Recall that, in our analysis:

1. total durables represent the *durables accumulated* over 2 consecutive periods. The distribution of total durables might then be expected to mimic the distribution of two-period incomes;

2. there are 2 channels of income generation – the labour market and the marriage market. Wealth begets wealth in the marriage market (through the signal function), but not in the labour market. However, income draws in the labour and marriage markets are *not independent* (as assumed in DGP4-5). In any period t,household decisions that boost labour market income (education) compete with decisions that boost marriage market income (durables);

3. we are interested in the steady state *equilibrium* distribution of income. This determines the susbet of income and durable levels that sustain each other in the long run.

Features (1) and (2) are model inputs – they represent the clustering mechanism embedded in the model assumptions (such as the signal function), and our notion of how income is generated (2 channels, dependent draws etc) that feeds the DGP. They shape the outcome (cluster formation) to an extent (as the exposition above demonstrates). However, the output of our model as represented by (3) above, provides economic meaning to the clusters and allows us to apply this insight to what is observed in empirical data.

To summarize:

1. Clustering is not caused by the simplified assumption of 2 wage levels in our model, as we show (above) that they are generated even when a continuous range of wages are allowed. Clustering occurs in *cumulated incomes* over 2 periods (thereby, cumulated durables over 2 periods), if high (low) incomes beget high (low) incomes over time.

2. The lowest cluster comprises households who are likely to have repeated draws of the lowest incomes *observed* over time. The goal of our paper is to demonstrate that the lowest cluster of households face a poverty of opportunities.

3. The number and shape of clusters is determined by "features" of the DGP over time – e.g. overlap of distributions by ϵ, δ, dependence between labour and marriage market draws etc – which are driven by model parameters and steady state dynamics. Clusters may be identified in nationally represented household data by an empirical procedure such as a mixture model.

5.2 Generalization: Allowing multiple wage levels $w_1, w_2, ..., w_n$ in the labour and marriage markets

Assumptions

1. In any period t, there are 3 generations in the households: Children, Parents and Grandparents.

2. Parents are the decision makers. Children and grandparents have no decision-making power but partake of household consumption in each period.

3. Given the draw of income I_t at the beginning of period t, parents in t choose their children's education level e_t ($e_t = e_H$ with cost $\widetilde{c}(e_H) = E$, or $e_t = e_L$ with cost $\widetilde{c}(e_L) = 0$). The residual income is spent on durables b_t ($I_t - C - E$,if $e_t = e_H$ is chosen, and $I_t - C$, if $e_t = e_L$ is chosen). C is the level of household subsistence consumption that must be met in any period.

4. Parents make decisions to maximize their expected lifetime utility of consumption: $U(c_t) + \delta E_t(c_{t+1})$, where $c_t = C + b_t + b_{t-1}$; $I_t = C + b_t + \widetilde{c}(e_t)$; $E_t c_{t+1} = E_t(I_{t+1} - \widetilde{c}(e_t) + b_t)$.

5. Durables last for 2 periods. Therefore, the total durables in the household at the beginning of period t is given by $(b_{t-1} + b_{t-2})$. b_t is chosen *during* period t,so at the beginning of period $(t+1)$, the total durables in the household is given by $(b_{t-1} + b_t)$.

6. Income drawn at the beginning of period t is the sum of wages earned in the labour and marriage markets, $w_1, w_2, ..., w_n$ where $w_1 < w_2 < ... < w_n$. The probability of earning w_i in the labour market is $\underline{p_i}(e)$ if the children of last period (and parents of the current period) turn out to have low productivity (w. p. q_L) and education level e invested in in the previous period; the probability of w_i is $\overline{p}_i(e)$ if productivity is high (w.p. $(1-q_L)$ and education level is e. The marriage market probability of earning w_i is $\widetilde{p}_i(B)$ where B is the total durables in the household at the beginning of the period. By assumption, $\widetilde{p}_n(B) = \Phi_S(\beta, \sigma^2)$, where β is the social standard for owning durables and σ^2 is popular sceptisim of the belief of β.

7. Example: Suppose, at the beginning of period t, parents draw income I_t and inherited durables b_{t-1}. Suppose parents choose $e_t = e_H$ during period t. This implies that durables b_t chosen in period t is $(I_t - C - E)$. Then, the expected income of parents in $(t+1)$ (who are children in t) is: $\sum_{i=1}^{n} [q_L \underline{p_i}(e_H) + (1-q_L)\overline{p}_i(e_H)].[\widetilde{p}_i(I_t - C - E + b_{t-1})](2w_i) + \sum_{i=1}^{n}\sum_{j>i}^{n} [\{q_L \underline{p_i}(e_H) + (1-q_L)\overline{p}_i(e_H)\}.\{\widetilde{p}_j(I_t - C - E + b_{t-1})\} + \{q_L \underline{p_j}(e_H) + (1-q_L)\overline{p}_j(e_H)\}.\{\widetilde{p}_i(I_t - C - E + b_{t-1})\}](w_i + w_j)$.

8. Since there are n possible wages in the labour and marriage market, there are $\frac{n(n+1)}{2}$ unique levels of total household income $(w_i + w_j)$ that may be drawn in any period. For each level of total income $(w_i + w_j)$, there are 2 possible levels of durables that may be chosen: $(w_i + w_j - C)$ (viz. when $e = 0$ is chosen) and $(w_i + w_j - C - E)$ (viz. when $e = 1$ is chosen). We assume that $C = 2w_1$, hence households with the lowest income $2w_1$ cannot afford high education $(e = e_H)$. Hence, there are $[n(n+1) - 1]$ levels of durables that may be observed in any period. The transition matrix that maps states (b_t, b_{t-1}) to (b_{t+1}, b_t) must therefore be of order $[n(n+1)-1]^2 \times [n(n+1)-1]^2$.

Optimization condition

As in the text, households will choose $e = e_L$ if the expected lifetime utility from choosing low education exceeds that of choosing high education. Given $I_t = w_k + w_l$ and inherited durables $b(e)$, say, a household's optimization condition can be written as follows:

$$\theta(w_k + w_l, b(e)) = 1 \text{ if } e = e_L, \text{ i.e. } Statement(1) \geq 0$$

$$\theta(w_k + w_l, b(e)) = 0 \text{ if } e = e_H, \text{i.e. } Statement(1) < 0$$

$$
\begin{aligned}
&Statement\ (1) : - \\
&E(1+\delta) \\
&+\delta T_L(q_L, \overline{p}_i(e), \underline{p_i}(e), w_1, w_2, ..., w_n, i = 1, 2, ..., n) \\
&+\delta T_M(\widetilde{p}_i(w_k + w_l - C - E + b(e)), \widetilde{p}_i(w_k + w_l - C + b(e)), w_1, w_2, ..., w_n) \\
&+\delta E T_E^1(q_L, \overline{p}_i(e), \underline{p_i}(e), w_1, w_2, ..., w_n, \widetilde{p}_i(w_i + w_j - C - E + b(e), i, j = 1, 2, ..., n) \\
&-\delta E T_E^0(q_L, \overline{p}_i(e), \underline{p_i}(e), w_1, w_2, ..., w_n, \widetilde{p}_i(w_i + w_j - C + b(e), i, j = 1, 2, ..., n)
\end{aligned}
$$

where,

$$T_L = q_L[\sum_{i=1}^{n} \{\underline{p_i}(e_L) - \underline{p_i}(e_H)\}w_i] + (1 - q_L)[\sum_{i=1}^{n} \{\overline{p}_i(e_L) - \overline{p}_i(e_H)\}w_i]$$

$$T_M = \sum_{i=}^{n} \{\widetilde{p}_i(w_k + w_l - C + b(e)) - \widetilde{p}_i(w_k + w_l - C - E + b(e))\}w_i\}$$

$$
\begin{aligned}
T_E^1 =& \sum_{i=1}^{n} [(1 - \theta(2w_i, w_k + w_l - C - E)].[\{q_L \underline{p_i}(e_H) + (1 - q_L)\overline{p}_i(e_H)\}\widetilde{p}_i(w_k + w_l - C - E + b(e))] \\
&+ \sum_{i=1}^{n}\sum_{j>i}^{n}\sum_{i=1}^{n} [(1 - \theta(w_i + w_j, w_k + w_l - C - E)].[\{q_L \underline{p_i}(e_H) + (1 - q_L)\overline{p}_i(e_H)\}\widetilde{p}_j(w_k + w_l - C - E + b(e)) + \{q_L \underline{p_j}(e_H) + (1 - q_L)\overline{p}_j(e_H)\}\widetilde{p}_i(w_k + w_l - C - E + b(e))]
\end{aligned}
$$

$$
\begin{aligned}
T_E^0 =& \sum_{i=1}^{n} [(1 - \theta(2w_i, w_k + w_l - C)].[\{q_L \underline{p_i}(e_L) + (1 - q_L)\overline{p}_i(e_L)\}\widetilde{p}_i(w_k + w_l - C + b(e))] \\
&+ \sum_{i=1}^{n}\sum_{j>i}^{n}\sum_{i=1}^{n} [(1 - \theta(w_i + w_j, w_k + w_l - C)].[\{q_L \underline{p_i}(e_L) + (1 - q_L)\overline{p}_i(e_L)\}\widetilde{p}_j(w_k + w_l - C + b(e)) + \{q_L \underline{p_j}(e_L) + (1 - q_L)\overline{p}_j(e_L)\}\widetilde{p}_i(w_k + w_l - C + b(e))]
\end{aligned}
$$

Notice that *Statement* (1) is of the form $f(w_k + w_l, b, \theta(I, b)/\{C, q_L, w_i, \overline{p}_i(.), \underline{p_i}(.), \widetilde{p}_i(.), i = 1, 2, ..., n\})$, for any k and l; where I represents all possible levels of income $(w_i + w_j)$ and b represents all possible levels of durables b. Hence the optimization condition can be written as:

$$\theta(w_k + w_l, b(e)) = 1 \text{ if } e = e_L$$

$$\theta(w_k + w_l, b(e)) = 0 \text{ if } e = e_H$$

i.e.,

$$\theta(w_k+w_l, b(e)) = 1 \text{ if } f(w_k+w_l, b, \theta(I,b)/\{C, q_L, w_i, \overline{p}_i(.), \underline{p_i}(.), \widetilde{p}_i(.), i = 1,2,...,n\}) \geq 0$$

$$\theta(w_k+w_l, b(e)) = 0 \text{ if } f(w_k+w_l, b, \theta(I,b)/\{C, q_L, w_i, \overline{p}_i(.), \underline{p_i}(.), \widetilde{p}_i(.), i = 1,2,...,n\}) < 0$$

The values of $\theta(I, b)$ are determined by the transition matrix in steady state. We compute the steady state transition matrix by an iterative procedure, given the assumed values of parameters. Here are the results from a model where $n = 3$, under the following parameter ranges:

$\boldsymbol{Set1}$: $w_1 \in (5, 20), w_2 \in (20, 30), w_3 \in (90, 120), p_1 \in (0, 0.15), p_2 \in (p_1, 0.25), p_3 \in (0.3, 0.5), p_4 \in (0.5, 0.7), p_5 \in (0.5, 0.7), p_6 \in (0.3, 0.5), p_7 \in (0.2, 0.25), p_8 \in (0, 0.15), \overline{\beta} \in (500, 3500), \sigma \in (100, 500), q_L = \delta = 0.5, \overline{C} = 2w_1, e = (0, emax), emax = w_1 + w_2 - \overline{C}$

Notice the additional assumptions we need to run the simulations for the parameters above (necessitated by the existence of more wage levels $(w_1, ..., w_n)$), viz.

1. The relative ranges of w_1, w_2 and w_3 (viz. how close or far away they are)
2. Labour market probabilities (we have to make assumptions about how education affects the probability of earning the middle wage w_2, versus only the assumption that high education increases the probability of earning the highest wage in the simplest model):
   (a) p_1 :probability of $w_3$when productivity is low and education is low
   (b) p_2 :probability of $w_3$when productivity is high and education is low
   (c) p_3 :probability of $w_3$when productivity is low and education is high
   (d) p_4 :probability of $w_3$when productivity is high and education is high
   (e) p_5 :probability of $w_1$when productivity is low and education is low

(f) p_6 :probability of $w_1$when productivity is high and education is low

(g) p_7 :probability of $w_1$when productivity is low and education is high

(h) p_8 :probability of $w_1$when productivity is high and education is high

3. Marriage market probabilities (we have to make assumptions about how more durables affects the probability of earning the middle wage w_2, versus only the assumption that B indicates a high wage in the marriage market with probability $\Phi_S(B/\overline{\beta}, \sigma)$):

(a) $\Phi_S(B/\overline{\beta}, \sigma)$:probability of earning $w_3$when total durables is $\overline{\beta}$

(b) $B[\frac{\Phi_S(B/\overline{\beta}, \sigma)}{1+B}]$:probability of earning $w_2$when total durables is $\overline{\beta}$

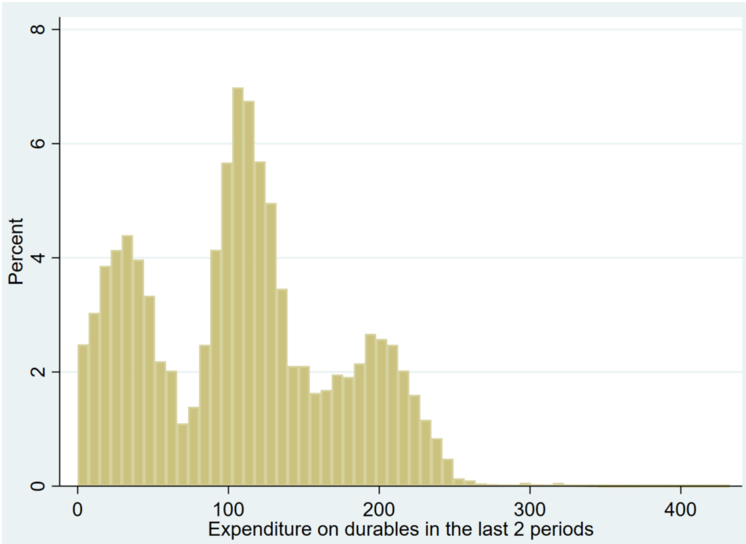
(c) $\frac{\Phi_S(B/\overline{\beta}, \sigma)}{1+B}$:probability of earning $w_1$when total durables is $\overline{\beta}$

4. Cost of education is such that only the lowest income group $(2w_1)$ cannot afford it.

[Note that (1)-(3) above amount to making a specific assumption about the wage distribution in the labour and marriage markets. We must be aware that these assumptions have an impact on the steady state clusters in and of themselves. For this reason, the extensions of the model make more sense when we have prior information about the assumptions to be made. These can then be imposed as calibrations of the parameters.]

The simulated dsitribution below represents 1000 households each, from 1000 communities (i.e. 1000 draws of parameters from uniform distributions over the ranges specified above).

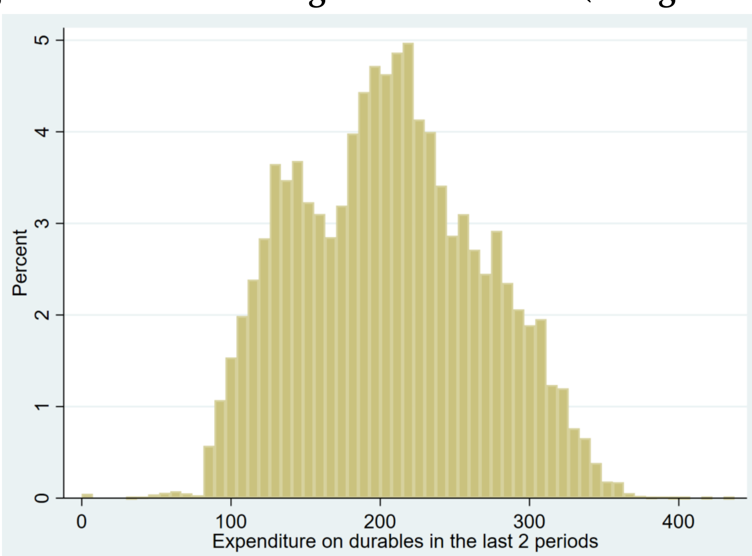
Fig. 11: Simulations using Set 1 Parameters (3 wages in labour and marriage markets)



Consider another example with the following set of parameters:

> ***Set2*** : $w_1 \in (5, 20), w_2 \in (60, 90), w_3 \in (90, 120), p_1 \in (0, 0.15), p_2 \in (p_1, 0.25), p_3 \in (0.3, 0.5), p_4 \in (0.5, 0.7), p_5 \in (0.5, 0.7), p_6 \in (0.3, 0.5), p_7 \in (0.2, 0.25), p_8 \in (0, 0.15), \overline{\beta} \in (500, 3500), \sigma \in (100, 500), q_L = \delta = 0.5, \overline{C} = 2w_1, e = (0, emax), emax = w_1 + w_2 - \overline{C}$

Fig 12: Simulations using Set 2 Parameters (3 wages in labour and marriage markets)



Recall that in Set 2, the range of the middle wage level w_2 is higher than in Set 1. This explains the disappearance of the lowest levels of 2-period expenditures ($\sim 0 - 75$) in steady state (Fig. 12).

5.3 Generalization: Allowing multiple education levels $e_1, e_2, ..., e_m$ with costs $\widetilde{c}(e_i)$

Assumptions

1. In any period t, there are 3 generations in the households: Children, Parents and Grandparents.

2. Parents are the decision makers. Children and grandparents have no decision-making power but partake of household consumption in each period.

3. Given the draw of income I_t at the beginning of period t, parents in t choose their children's education level e_t, where $e_t \in \{e_1, e_2, ..., e_m\}$ with cost $\widetilde{c}(e_1) = 0$). The residual income is spent on durables b_t. C is the level of household subsistence consumption that must be met in any period. Assume $0 = \widetilde{c}(e_1) < \widetilde{c}(e_2) < ... < \widetilde{c}(e_m)$;

4. Parents make decisions to maximize their expected lifetime utility of somsumption: $U(c_t) + \delta E_t(c_{t+1})$, where $c_t = C + b_t + b_{t-1}$; $I_t = C + b_t + \widetilde{c}(e_t)$; $E_t c_{t+1} = E_t(I_{t+1} - \widetilde{c}(e_t) + b_t)$.

5. Durables last for 2 periods. Therefore, the total durables in the household at the beginning of period t is given by $(b_{t-1} + b_{t-2})$. b_t is chosen *during* period t,so at the beginning of period $(t+1)$, the total durables in the household is given by $(b_{t-1} + b_t)$.

6. Income drawn at the beginning of period t is the sum of wages earned in the labour and marriage market, w_1, w_2 where $w_1 < w_2$. The probability of earning w_i in the labour market is $\underline{p_i}(e)$ if the children of last period (and parents of the current period) turn out to have low productivity (w. p. q_L) and education level e invested in in the previous period; the probability of w_i is $\overline{p}_i(e)$ if productivity is high (w.p. $(1 - q_L)$ and education level is e. Assume: $\underline{p_2}(e_i) < \underline{p_2}(e_j)$ if $\widetilde{c}(e_i) < \widetilde{c}(e_j)$; $\overline{p}_2(e_i) < \overline{p}_2(e_j)$ if $\widetilde{c}(e_i) < \widetilde{c}(e_j)$ and $\underline{p_1}(e_i) > \underline{p_1}(e_j)$ if $\widetilde{c}(e_i) < \widetilde{c}(e_j)$; $\overline{p}_1(e_i) < \overline{p}_1(e_j)$ if $\widetilde{c}(e_i) < \widetilde{c}(e_j)$. These assumptions ensure that more costly education increases (decreases) the probability of the highest (lowest) wage possible in the labour market.

7. The marriage market probability of earning w_i is $\widetilde{p}_i(B)$ where B is the total durables in the household at the beginning of the period. By assumption, $\widetilde{p}_n(B) = \Phi_S(\beta, \sigma^2)$, where β is the social standard for owning durables and σ^2 is popular scepticism of the belief of β.

8. Example: Suppose, at the beginning of period t, parents draw income I_t and inherited durables b_{t-1}. Suppose parents choose $e_t = e_k$ during period t. This implies that durables b_t chosen in period t is $(I_t - C - \widetilde{c}(e_k))$. Then, the expected income of parents in $(t+1)$ (who are children in t) is: $\sum_{i=1}^{n} [q_L \underline{p_i}(e_k) + (1-q_L)\overline{p}_i(e_k)].[\widetilde{p}_i(I_t - C - \widetilde{c}(e_k) + b_{t-1})](2w_i) + \sum_{i=1}^{n}\sum_{j>i}^{n} [\{q_L \underline{p_i}(e_k) + (1-q_L)\overline{p}_i(e_k)\}.\{\widetilde{p}_j(I_t - C - \widetilde{c}(e_k) + b_{t-1})\} + \{q_L \underline{p_j}(e_k) + (1-q_L)\overline{p}_j(e_k)\}.\{\widetilde{p}_i(I_t - C - \widetilde{c}(e_k) + b_{t-1})\}](w_i + w_j)$.

9. Since there are 2 possible wages in the labour and marriage market, there are 3 unique levels of total household income $(w_i + w_j)$ that may be drawn in any period. For each level of total income $(w_i + w_j)$, there are m possible levels of durables that may be chosen: $(w_i + w_j - \widetilde{c}(e_k))$ $(k = 1, 2, ..., m)$. We assume that $C = 2w_1$, hence households with the lowest income $2w_1$ can only afford education e_1 with cost 0. Hence, there are $(2m+1)$ levels of durables that may be observed in any period. The transition matrix that maps states (b_t, b_{t-1}) to (b_{t+1}, b_t) must therefore be of order $[2m+1]^2 \times [2m+1]^2$.

Optimization condition

Suppose income drawn is $I = w_k + w_l$ and inherited durables is b_{t-1}.

$Exp(e = e_i)$ =Expected lifetime consumption if $e_t = \widetilde{c}(e_i)$ is chosen $= C + b_{t-1} + (I - C - \widetilde{c}(e_i)) + \delta E_t[C + (I - C - \widetilde{c}(e_i)) + b_{t+1}] = b_{t-1} + I(1+\delta) - (1+\delta)\widetilde{c}(e_i) + \delta E_t[I_{t+1} - C - \widetilde{c}(e_{t+1})/e_t = e_i]$

$Exp(e = e_{i+1})$ =Expected lifetime consumption if $e_t = \widetilde{c}(e_{i+1})$ is chosen $= C + b_{t-1} + (I - C - \widetilde{c}(e_{i+1})) + \delta E_t[C + (I - C - \widetilde{c}(e_{i+1})) + b_{t+1}] = b_{t-1} + I(1+\delta) - (1+\delta)\widetilde{c}(e_{i+1}) + \delta E_t[I_{t+1} - C - \widetilde{c}(e_{t+1})/e_t = e_{i+1}]$

Therefore,

$Exp(e = e_i) - Exp(e = e_{i+1}) =$

$b_{t-1} + I(1+\delta) - (1+\delta)\widetilde{c}(e_i) + \delta E_t[I_{t+1} - C - \widetilde{c}(e_{t+1})/e_t = e_i]$

$-b_{t-1} - I(1+\delta) + (1+\delta)\widetilde{c}(e_{i+1}) - \delta E_t[I_{t+1} - C - \widetilde{c}(e_{t+1})/e_t = e_{i+1}]$

$= (1+\delta)[\widetilde{c}(e_{i+1}) - \widetilde{c}(e_i)] + \delta E_t[I_{t+1}/e_t = e_i] - \delta E_t[\{I_{t+1}/e_t = e_{i+1}] - E_t[\widetilde{c}(e_{t+1})/e_t = e_i] + \delta E_t[\widetilde{c}(e_{t+1})/e_t = e_{i+1}]$

Let T_1 $=(1+\delta)[\widetilde{c}(e_{i+1}) - \widetilde{c}(e_i)]$

Let $T_2 = E_t[I_{t+1}/e_t = e_i]$

Let T_3 $=E_t[\{I_{t+1}/e_t = e_{i+1}]$

Let T_4 $=E_t[\widetilde{c}(e_{t+1})/e_t = e_i]$

$Let T_5 = E_t[\widetilde{c}(e_{t+1})/e_t = e_{i+1}]$

Look at T_2, given income $(w_k + w_l)$ and inherited durables b_{t-1}:

$T_2(e_i) = [q_L \underline{p}_1(e_i) + (1-q_L)\overline{p}_1(e_i)].[\widetilde{p}_1((w_k+w_l) - C - \widetilde{c}(e_i) + b_{t-1})](2w_1) + [q_L \underline{p}_2(e_i) + (1-q_L)\overline{p}_2(e_i)].[\widetilde{p}_2((w_k+w_l) - C - \widetilde{c}(e_i) + b_{t-1})](2w_2) + [\{q_L \underline{p}_1(e_i) + (1-q_L)\overline{p}_1(e_i)\}.\{\widetilde{p}_2((w_k + w_l) - C - \widetilde{c}(e_i) + b_{t-1})\} + \{q_L \underline{p}_2(e_i) + (1-q_L)\overline{p}_2(e_i)\}.\{\widetilde{p}_1((w_k+w_l) - C - \widetilde{c}(e_i) + b_{t-1})\}](w_1 + w_2)]$

Look at T_3:

$T_3(e_{i+1}) = [q_L \underline{p}_1(e_{i+1}) + (1-q_L)\overline{p}_1(e_{i+1})].[\widetilde{p}_1((w_k+w_l) - C - \widetilde{c}(e_{i+1}) + b_{t-1})](2w_1) + [q_L \underline{p}_2(e_{i+1}) + (1-q_L)\overline{p}_2(e_{i+1})].[\widetilde{p}_2((w_k+w_l) - C - \widetilde{c}(e_{i+1}) + b_{t-1})](2w_2) + [\{q_L \underline{p}_1(e_{i+1}) + (1-q_L)\overline{p}_1(e_{i+1})\}.\{\widetilde{p}_2((w_k + w_l) - C - \widetilde{c}(e_{i+1}) + b_{t-1})\} + \{q_L \underline{p}_2(e_i) + (1-q_L)\overline{p}_2(e_i)\}.\{\widetilde{p}_1((w_k + w_l) - C - \widetilde{c}(e_{i+1}) + b_{t-1})\}](w_1 + w_2)]$

Below, we present the simulation results for 2 sets of parameters (Set 3 and Set 4). The parameters that need to be specified to run these simuations are as follows:

1. Low and high wage in the labour/marriage markets, w_L, w_H

2. Labour market probabilities (we have to make assumptions about how different levels of education affects the probability of low and high wage):

   (a) p_1 :probability of w_H when productivity is low and education is low ($e_1 = 0$)

   (b) p_2 :probability of w_H when productivity is high and education is low ($e_1 = 0$)

   (c) p_3 :probability of w_H when productivity is low and education is medium (e_2)

   (d) p_4 :probability of w_H when productivity is high and education is medium (e_2)

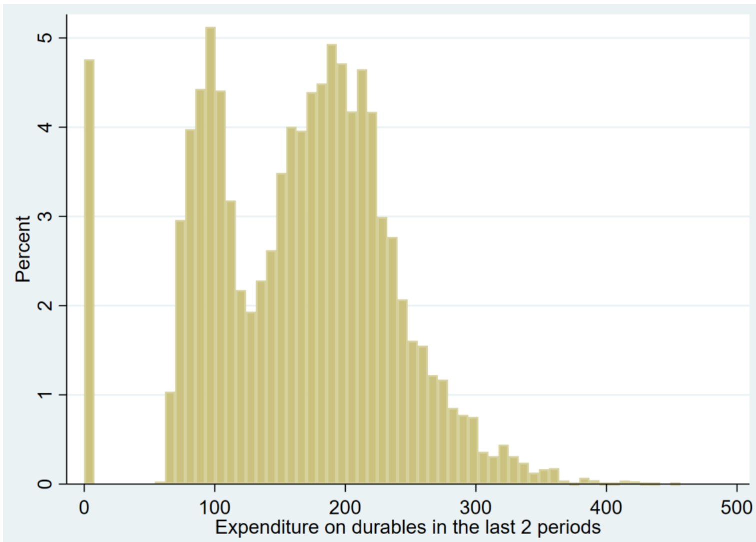
   (e) p_5 :probability of w_H when productivity is low and education is high (e_3)

   (f) p_6 :probability of w_H when productivity is high and education is high (e_3)

3. The costs of medium and high education (e_2 and e_3) are assumed to be such that only the lowest income group (who earn income $2w_L$) cannot afford it.

$\boldsymbol{Set3}$: $w_L \in (5, 20), w_H \in (80, 120), p_1 \in (0, 0.15), p_2 \in (p_1, 0.35), p_3 \in (0.35, 0.5), p_4 \in (p_3, 0.75), p_5 \in (0.55, 0.7), p_6 \in (p_5, 0.95), \overline{\beta} \in (500, 3500), \sigma \in (100, 500), q_L = \delta = 0.5, \overline{C} = 2w_L, e_1 = 0, e_2 = (0, w_L + w_H - \overline{C}), e_3 = (e_2, w_L + w_H - \overline{C})$

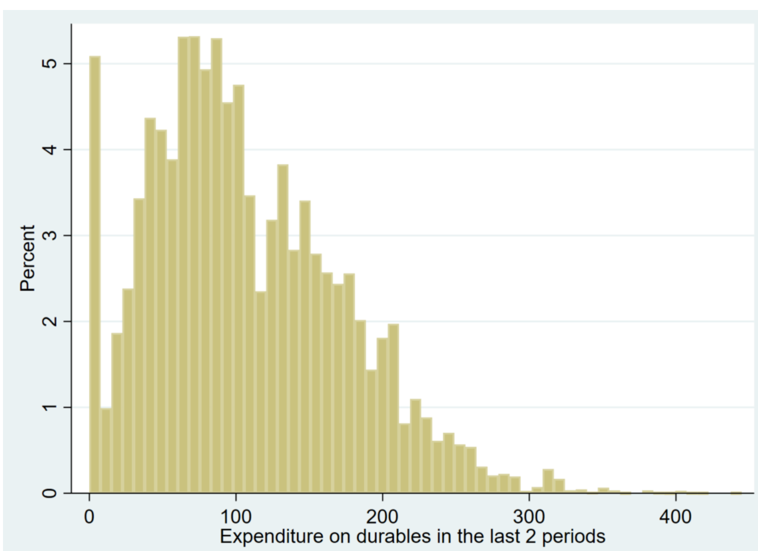
Figure 13: Simulations using Set 3 Parameters (3 levels of education in the labour market)



***Set*4** : $w_L \in (5,80), w_H \in (80,120), p_1 \in (0,0.15), p_2 \in (p_1,0.35), p_3 \in (0.35,0.5), p_4 \in (p_3,0.75), p_5 \in (0.55,0.7), p_6 \in (p_5,0.95), \overline{\beta} \in (500,3500), \sigma \in (100,500), q_L = \delta = 0.5, \overline{C} = 2w_L, e_1 = 0, e_2 = (0, w_L + w_H - \overline{C}), e_3 = (e_2, w_L + w_H - \overline{C})$

(Same parameter ranges as Set 3, except that the ranges of $w_L$and w_H are abutting.)

Figure 14: Simulations using Set 4 Parameters (3 levels of education in the labour market)



As in the simulation presented in the manuscript, when the range of $w_H$and w_L are contiguous, the clusters at positive levels of 2-period expenditure move to the left.

5.4 Further extensions of the model

- Incorporate more "opportunities" (in addition to labour market and marriage market opportunities)

There are 2 broad ways to think about expanding the range of "opportunities" available to households:

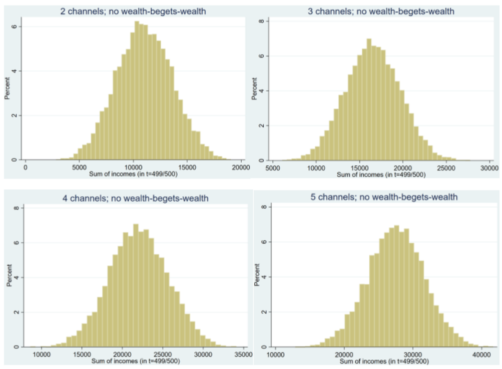
1. by increasing the number of channels in which households may earn income. Savings/investment opportunities, or opportunities in a foreign labour market (see the discussion of "networks" in Munshi (2014)) could be modelled in this manner, viz. as a third channel of income generation. (Networks based on social identity could also be modelled as segmentation in the domestic labour and marriage markets; here social identity, say ω, would play a role alongside education (durables) to determine labour (marriage) market matching probabilities).

2. by imposing additional features that indirectly lead to higher income generation in the original channels (labour and marriage markets). Take access to credit, for example. In rural India, durable goods often provide collateral value for procuring informal loans. This effect could be incorporated in the model by introducing an interest parameter r that is charged on loans obtained using durable goods as collateral. The availability of such loans would ease period-wise liquidity-constraints so expenditure could exceed current income by the amount of the loan taken. However, the extent to which this is possible would depend on the level of durables already owned by the household (since this is what determines loan availability). Thus, the availability of credit (hence, the abillity to spend more on future-income-enhancing expenses such as education/ durables) – to households *already* owning durables – provides yet another boost to the wealth-begets-wealth mechanism in the model. Households that *do not* own durables are even more likely to fall behind in the lowest cluster (compared with the case where there is no access to credit, as in the current model). Alternatively, we could interpret the lowest cluster of households as lacking access to 3 sources of opportunities: the labour market, the marriage market and the credit market.

To further understand the role of (1) above (viz. adding more channels of income generation to the model) we conduct the experiments below.

Let us focus again on step (B) of Fig. A. Consider the distribution of 2-period incomes that would be obtained if there were 2, 3, 4, and 5 channels of income generation in an economy, but with no wealth-begets-wealth mechanism. In the examples below, we have income from each channel drawn independently from $U(500{,}5000)$; and the total income in any period is the sum of incomes earned in the number of channels that exist in the economy.

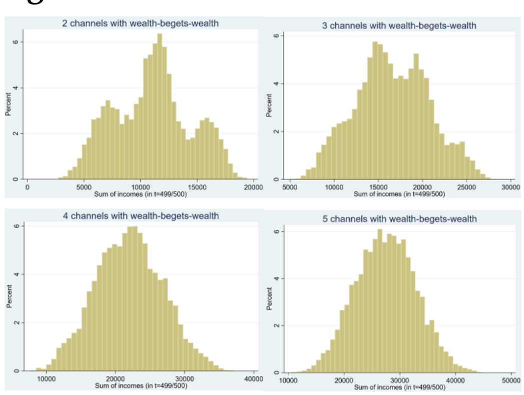
Figure 15.1



Notice that adding channels does not change the nature of two-period incomes when there is no wealth-begets-wealth mechanism, mirroring our expectation that the sum of independent uniform distributions approach a normal distribution.

Now consider adding channels that have an embedded wealth-begets-wealth mechanism. In the examples below, income in each channel is drawn from $U(500, 3000)$ if income in the previous period (in this channel) is below 3000; else income is drawn from $U(3000, 5000)$ (independent draws across channels). Total household income in any period is the sum of incomes earned in the n channels.

Figure 15.2



It is immediately evident that as more and more channels are added (independently), the clusters progressively dissolve, even when there is an explicit wealth-begets-wealth mechanism at play in each channel. Thus the clustering impact of a wealth-begets-wealth mechanism "bites" when (a) opportunites (embodied by the number of channels of income generation) are limited; and, *most importantly*, (2) when scarcity of household resources restrict access to the different channels. In other words, resource scarcity (such as income that must be shared between education and durable spending) introduces a *de-*

pendence in the income draws that a household makes from different channels (violating the *independence* assumption reflected in Figures 15.1-15.2). The economic model in Section 2 drives the nature of dependence in income draws across the available channels of "opportunites".

- Incorporate inter-generational effects of education through the productivity parameter

There is empirical evidence that parental education has positive effects on offspring such as in child health and productivity outcomes. Effects such as these may be incorporated into the model through the productivity parameter α_L. In the basic model, children draw a low productivity level α_L with probability q_L in every period. However, we could postulate that q_L depends on parental education (i.e. more educated parents have a lower probability of bearing offspring with low productivity). An interesting experiment would be to compare the outcomes in the case where parents are aware of (or believe in) the positive intergenerational effect on productivity versus the case where parents ignore the intergenerational effect in their decision-making.

- Group dynamics, network effects

Group dynamics could be introduced in the model (e.g. caste and other exclusionary politics) by allowing group identity, say ω, to determine the social standard β and the high-wage earning probability in the labour market. This approach could also be used to analyze network effects in labour and marriage market matching (Munshi (2014)). In this approach, group identity would constitute an easily-observed (pre-determined) characteristic of households. (How is the social standard β *determined* in steady state (the "caste system" in any society) given a certain distribution of income and group characteristics? A general equilibrium framework – see below – with political (e.g. majority) "voting" on the social standard, may offer insights into this question.)

- General equilibrium analysis by linking to the macroeconomy (parameters become endogenous; initial conditions will start to matter)

Embedding an economy-wide production function (with unkilled labour, skilled labour) in the model would permit the endogenization of the wage parameters of the model. Embedding a population growth process could enable an endogenous matching mechanism, of households to jobs in the labour market (driven by technology) and households to each other in the marriage market – endogenizing the probabilities of high wage in the labour and marriage markets. An endogenous matching process within a general equilibrium framework would be especially interesting for examining the impact of network effects (suggested by Munshi (2014)) on long-run growth and development.