Abstract

In this paper we develop a theory for the evolution of the American dream and explore its economic and social consequences. We model the American dream and its change over time as the outcome of a learning model in which individuals form beliefs about the return to human capital investment from the experience of their neighbors. Increasing inequality drives up segregation across communities in the U.S. and reduces interactions among different socio-economic groups. In turn, less interaction translates into fewer shared success stories. As lower income Americans are increasingly surrounded by poverty and failure, they shy away from the American dream and invest less in their children than higher income parents. This gap widens over time reducing intergenerational mobility and feeding higher levels of inequality in the following generations. To investigate the empirical relevance of our theory, we use a new county-level data set to compare our calibrated model to the time-series and geographic patterns of education.

*The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.
1 Introduction

That hard work and talent lead to success regardless of one’s background has long been a central tenet of American culture. Recently, however, the belief in the American Dream seems to have weakened, as surveys show that low- and middle-income Americans are increasingly pessimistic about their children’s prospects. On the surface, this increased pessimism appears to be fueled by the significant increase in inequality that has taken place over the last forty years. However, inequality in outcomes needn’t be problematic, as long as such dispersed outcomes are driven by individual merit to a large degree. Instead, we argue that another force is driving the apparent twilight of the American Dream: income-based segregation across communities in the U.S. means less and less interactions among the different socio-economic classes. In turn, less interaction means fewer shared success stories. As lower income Americans are increasingly surrounded by poverty and failure, they start to doubt the American Dream. These doubts become self-fulfilling, as the resulting underinvestment in human capital limits the upward mobility of the subsequent generation.

We model the American Dream and its evolution over time as the outcome of a learning model in which the return to human capital evolves stochastically. Individuals invest in the human capital of their children based on their beliefs about this return. The nature of the learning process is local: Individuals live in different communities and learn both from their own experiences and from the experiences of their neighbors. Individuals are heterogeneous along ability and educational attainment and communities differ in the distribution of these characteristics. More segregated communities exhibit more uniformity along the educational attainment dimension. The local nature of learning implies that segregation limits the pool of outcomes observed. Since learning only takes place when observing the outcomes of individuals who benefit from prior investment, communities with low investment rates are characterized by slow learning and as a result experience less upward mobility compared with communities with higher initial investment rates. As the skill premium grows, segregated communities form diverging beliefs about its evolution and grow further apart in terms of education and income levels. In a version of the model where segregation is the endogenous outcome of residential choices of individuals, an increase in inequality fuels further
segregation, a weakening of the American Dream, and further divergence in outcomes for subsequent generations. We discipline the theory with empirical evidence on the evolution of inequality and segregation over the last forty years.

We use the theory to analyze different types of policies. Policies geared towards reducing current income inequality are potentially ineffective unless they also change people’s beliefs about the returns to investing in human capital.

2 Empirics

To create our main dataset, we used the tables downloaded from Social Explorer. This dataset contains socio-demographic data on the United States from various surveys, including the decennial census. Using census tract level data, we generated MSA-level Gini coefficients, segregation indexes, and county-level spatial correlation measures for 1970-2000.

There is a large literature on measures of segregation. Massey and Denton (1988) grouped the measures into 5 key dimensions: evenness, exposure, concentration, centralization, and clustering. Within each dimension there can be several measures with different interpretations. Segregation indexes are interpreted as the separation of one group (black), called the “minority”, the from the “majority” (white). For racial segregation results, we assign black to be the minority group and non-Hispanic whites to be the majority. In our calculations for the segregation of the educated, we define the minority population as those with a college degree or more, and the complement of this set (those with a high school diploma or less) as the majority.

The dimension we chose to focus on is exposure; the measure is isolation. The formula we used is from “Appendix B: Measures of Residential Segregation” published by the U.S. Census Bureau. Segregation indexes use a geographic unit (country, state, county, MSA, etc) and a subunit (census tract, city block, etc). In the segregation indexes we calculated, the unit was MSA, and the subunit was census tract.

The isolation index measures “the extent to which minority members are exposed only to one another” (Massey and Denton, 1988). It is the minority-weighted average of the minority proportion in each area.
The formula for the isolation index in MSA \( j \) is:

\[
I_j = \sum_{i=1}^{n} \left( \frac{x_i}{X} \right) \left( \frac{x_i}{t_i} \right)
\]

Here \( t_i \) is the total population in census tract \( i \), \( x_i \) is the minority population of census tract \( i \), and \( X \) is the total minority population of MSA \( j \).

Moran’s I is a measure of spatial correlation (Moran 1950). It is a measure of local geographic similarity for a variable \( \epsilon \) commonly used in fields such as geography, sociology, and epidemiology to measure spatial effects.

The formula for Moran’s I in county \( j \) is:

\[
MI_j = \left( \frac{N}{\sum_i \sum_k \omega_{i,k}} \right) \frac{\sum_i \sum_j \omega_{i,k} \epsilon_i \epsilon_k}{\sum_i \epsilon_i^2}
\]

Where \( \epsilon_i \) is the variable value in census tract \( i \), \( N \) is the number of census tracts in county \( j \), and \( \omega_{i,k} \) is a weighting matrix that gives higher weights to census tracts that are closer together.

Although racial segregation looms large in popular consciousness, Figure 1 shows that the racial isolation index has decreased from 1970 to 2000.

![Figure 1: MSA Level Mean: Racial Isolation](image)

By contrast, Figure 2 shows how segregation of the educated has increased over time.
Over this time period, the Gini coefficient measured at the MSA-level has increased (Figure 3).

Finally, the correlation between inequality as measured by the Gini coefficient and isolation of the educated has been increasing over the same time period, as shown in Figure 4.
In addition, the county-level Moran’s I shows an increasing time trend (Figure 5). This suggests that the proportion of the population that has a college education is increasingly spatially correlated (clustered in space).

3 A Basic Model with Exogenous Segregation

We consider an overlapping generations economy with agents indexed by $i$. Agents live for two periods: they accumulate human capital when young, and they work, consume and raise
one child when old. Parents choose whether or not to invest in the human capital of their children, \( h_{it} \in \{0, 1\} \). This choice affects the parent’s consumption today and expectations of the child’s wage tomorrow, both of which enter the parent’s utility:

\[
U_{it} = \frac{c_{it}^{1-\gamma}}{1-\gamma} + \beta E_{it} \left( \frac{w_{it+1}^{1-\gamma}}{1-\gamma} \right), \quad \gamma > 1, \quad \beta \in (0, 1).
\]  

(1)

For simplicity, agents supply one unit of labor inelastically when working. The budget constraint of agent \( i \) born in period \( t-1 \) is

\[
c_{it} = w_{it} - \tau h_{it},
\]

(2)

where \( \tau > 0 \) is the cost of education. The effect of education on the child’s future wage is given by

\[
w_{it+1} = \exp \{ a_{it+1} + \theta h_{it} \},
\]

(3)

where \( a_{it+1} \sim N(\mu_a, \sigma_a^2) \) is the child’s unobserved ability and \( \theta > 0 \) represents the value of education, assumed to be independent of ability.

(i) Only parents with wages larger than the fixed education cost, \( w_{it} > \tau \), will even contemplate education.

(ii) The parent makes the education choice without knowing the child’s ability.

The parent’s problem becomes

\[
\max_{h_{it} \in \{0,1\}} U_{it} = \frac{(w_{it} - \tau h_{it})^{1-\gamma}}{1-\gamma} + \frac{\beta}{1-\gamma} E_{it} \left[ \exp \left\{ (1-\gamma) (a_{it+1} + \theta h_{it}) \right\} \right],
\]

where expectations are taken over the distribution of ability and of returns to education, which are assumed to be independent. Hence

\[
U_{it} = \frac{(w_{it} - \tau h_{it})^{1-\gamma}}{1-\gamma} + \frac{\beta}{1-\gamma} \exp \left\{ (1-\gamma) (\mu_a + h_{it} \mu_{it}) + \frac{(1-\gamma)^2}{2} (\sigma_a^2 + h_{it} \sigma_{it}^2) \right\},
\]
where
\[ \mu_{it} \equiv E_{it} [\theta] \quad \text{and} \quad \sigma_{it}^2 \equiv V_{it} [\theta]. \] (4)

Let
\[ M_a \equiv \exp \left\{ (1 - \gamma) \mu_a + \frac{(1 - \gamma)^2 \sigma_a^2}{2} \right\}. \] (5)

Then
\[ U_{it} = \frac{(w_{it} - \tau h_{it})^{1-\gamma}}{1 - \gamma} + \frac{\beta M_a}{1 - \gamma} \exp \left\{ (1 - \gamma) h_{it} \mu_{it} + \frac{(1 - \gamma)^2}{2} h_{it} \sigma_{it}^2 \right\}. \] (6)

The agent compares the utility when investing with the utility when not investing. The utility when not investing is
\[ U_{it}^0 = \frac{w_{it}^{1-\gamma} + \beta M_a}{1 - \gamma}. \] (7)

The utility when investing is
\[ U_{it}^1 = \frac{(w_{it} - \tau)^{1-\gamma}}{1 - \gamma} + \frac{\beta M_a}{1 - \gamma} \exp \left\{ (1 - \gamma) \mu_{it} + \frac{(1 - \gamma)^2}{2} \sigma_{it}^2 \right\}. \] (8)

If \( \theta \) is known: The utility when investing is
\[ U_{it}^1 = \frac{(w_{it} - \tau)^{1-\gamma}}{1 - \gamma} + \frac{\beta M_a}{1 - \gamma} \exp \{(1 - \gamma) \theta\}. \] (9)

The difference between \( U_{it}^1 \) and \( U_{it}^0 \) determines the threshold for investment. The parent invests in the child’s education iff
\[ \beta M_a [1 - \exp \{(1 - \gamma) \theta\}] \geq (w_{it} - \tau)^{1-\gamma} - w_{it}^{1-\gamma}. \] (10)

Comparative statics:

(i) Investment in education is decreasing in the average ability in the economy, since higher expected ability implies higher expected wages for offsprings (\( M_a \) is decreasing in \( \mu_a \)). This result arises because we assume that education and ability are not complementary.

(ii) Conversely, a higher dispersion in ability increases investment, since it triggers a precautionary motive (\( M_a \) is increasing in \( \sigma_a^2 \)).
(iii) Investment in education is decreasing in its cost (the RHS of the inequality is increasing in $\tau$).

**Beliefs:** Suppose that $\theta$ is not known. Agent $i$’s beliefs at time $t$ about the value of $\theta$ are $\hat{\theta}_{it} \sim \mathcal{N}(\mu_{it}, \sigma^2_{it})$. Then, the utility when investing is

$$U^1_{it} = \frac{(w_{it} - \tau)^{1-\gamma}}{1-\gamma} + \beta M_a \exp\left\{ (1-\gamma) \mu_{it} + \frac{(1-\gamma)^2}{2} \sigma^2_{it} \right\}.$$  \hspace{1cm} (11)

The parent invests in the child’s education iff

$$\beta M_a \left[ 1 - \exp\left\{ (1-\gamma) \mu_{it} + \frac{(1-\gamma)^2}{2} \sigma^2_{it} \right\} \right] \geq (w_{it} - \tau)^{1-\gamma} - w_{it}^{1-\gamma}. \hspace{1cm} (12)$$

**Comparative statics:**

(i) More optimistic beliefs about the return to education lead to more investment (the LHS of the inequality is increasing in $\mu_{it}$).

(ii) More uncertainty about the return to education leads to less investment (the LHS of the inequality is decreasing in $\sigma^2_{it}$).

(iii) As in the case of full information about $\theta$, investment in education is higher for lower average ability, for higher dispersion in ability, and for lower costs.

**Learning:** At time $t = 0$, all agents have beliefs about the return to education drawn from the same distribution, which is given exogenously: $\hat{\theta}_{i0} \sim \mathcal{N}(\mu_0, \sigma^2_0)$. Hence, parents in the first generation only differ in their investment decisions because their wages differ. In turn, wages differ because abilities differ.

Subsequent generations inherit the beliefs of their parents and update these beliefs using their parents’ investment decision and their own wage: $h_{it-1}$ and $w_{it}$ (under the assumption that they do not know their own ability). They also draw additional signals from their peers: they observe the wages and education levels of $J - 1$ peers.

Let $J_i$ denote the set of signals observed by agent $i$ (including the agent’s own wage and education level). Let $\hat{\mu}_{it}$ and $\hat{\sigma}^2_{it}$ denote the mean and variance of the estimator $\hat{\theta}$ obtained
from running the regression

$$\log w_{jt} = \mu_a + \theta h_{jt-1} + \varepsilon_{jt}, \quad \text{for } j \in J_i,$$

(13)

where $\varepsilon_{jt} \equiv a_{jt} - \mu_a$. Let $h_{it-1} \equiv \sum_{j \in J_i} h_{jt-1}$. The estimated coefficient is normally distributed with mean and variance

$$\begin{align*}
\tilde{\mu}_{it} = & \frac{1}{h_{it-1}} \sum_{j \in J_i} (\log w_{jt} - \mu_a) h_{jt-1}, \\
\tilde{\sigma}_{it}^2 = & \sigma_a^2 h_{it-1}.
\end{align*}$$

(14)

Agent $i$’s posterior beliefs about the return to education are then normally distributed with mean and variance

$$\begin{align*}
\mu_{it} = & \frac{\tilde{\sigma}_{it}^2}{\sigma_a^2 + \sigma_{\eta_{it}}^2} \mu_{it-1} + \frac{\sigma_{\eta_{it}}^2}{\sigma_a^2 + \sigma_{\eta_{it}}^2} \tilde{\mu}_{it}, \\
\sigma_{it}^2 = & \frac{\tilde{\sigma}_{it}^2 + \sigma_{\eta_{it}}^2}{\sigma_a^2 + \sigma_{\eta_{it}}^2}.
\end{align*}$$

(15)

**Learning with Noise:** Suppose that agents observe their peers’ wages with some noise: for $j \in J_i$, the observed wage of peer $j$ in period $t$ is

$$\log w^o_{jt} = \log w_{jt} + \eta_{jt}, \quad \eta_{jt} \sim \mathcal{N}(0, \sigma_{\eta_{it}}^2),$$

(16)

where $\sigma_{\eta_{it}}^2$ may be time-varying (to keep up with the dispersion in actual wages over time). The wage regression becomes

$$\log w^o_{jt} = \mu_a + \theta h_{jt-1} + \varepsilon^o_{jt}, \quad \text{for } j \in J_i,$$

(17)

where $\varepsilon^o_{jt} \equiv a_{jt} - \mu_a + \eta_{jt}$. As above, let $h_{it-1} \equiv \sum_{j \in J_i} h_{jt-1}$. The estimated coefficient is now normally distributed with mean and variance

$$\begin{align*}
\tilde{\mu}_{it} = & \frac{1}{h_{it-1}} \sum_{j \in J_i} (\log w^o_{jt} - \mu_a) h_{jt-1}, \\
\tilde{\sigma}_{it}^2 = & \frac{\sigma_a^2 + \sigma_{\eta_{it}}^2}{h_{it-1}}.
\end{align*}$$

(18)
Since the agent’s observations are now less precise, the agent places less weight on learning from peers, and hence learning slows down.

### 3.1 Numerical Results

Consider the parameterization shown in Table I and the results shown in Table II.

**Table I: Parameter Values**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Explanation/Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount factor</td>
<td>$\beta$</td>
<td>0.9</td>
<td>Standard</td>
</tr>
<tr>
<td>Risk aversion</td>
<td>$\gamma$</td>
<td>3</td>
<td>Standard</td>
</tr>
<tr>
<td>Ability mean</td>
<td>$\mu_a$</td>
<td>14</td>
<td>Avg lifetime earnings in 1970 = 1.5 million (AWI)</td>
</tr>
<tr>
<td>Ability std dev</td>
<td>$\sigma_a$</td>
<td>0.4</td>
<td>90-10 wage diff in 1970 = 1.15 (AKK 08)</td>
</tr>
<tr>
<td>Value of education</td>
<td>$\theta$</td>
<td>0.6</td>
<td>Avg log college premium = 0.55 (AKK 08)</td>
</tr>
<tr>
<td>Cost of education</td>
<td>$\tau$</td>
<td>40000</td>
<td>Avg 4y college cost to lifetime earnings = 0.03 (Census)</td>
</tr>
<tr>
<td>Initial beliefs mean</td>
<td>$\mu_0$</td>
<td>0.6</td>
<td>Unbiased initial beliefs</td>
</tr>
<tr>
<td>Initial beliefs std dev</td>
<td>$\sigma_0$</td>
<td>0.768</td>
<td>College completion rate in 1970 = 0.14 (College Board)</td>
</tr>
<tr>
<td>Number of peers</td>
<td>$n_P$</td>
<td>4</td>
<td>College completion rate in 2010 = 0.3 (College Board)</td>
</tr>
</tbody>
</table>

The discount factor and the risk aversion coefficient are set to standard values in the literature. The mean and standard deviation of ability are set to generate the initial average real lifetime earnings of $1.5$ million in 1970 (in 2014 dollars), and the initial 90-10 log wage differential of 1.15 in 1970 (Autor, Katz & Kearney (2008)). The value of education is set to match the average log college/high school wage gap of roughly 0.55 over the period 1970-2005 (Autor et al. (2008)). The cost of education is set to match the average cost of college (tuition, fees, room and board, net of grant aid and tax benefits) relative to average lifetime earnings, of roughly 0.03 (College Board). The mean of initial beliefs is set to the value of education (unbiased beliefs), while the standard deviation of beliefs is set to target the initial college completion rate of 14% in 1970 (Census). Finally, the number of peers individuals learn from generates different speeds of growth in college attainment, which in the data had doubled to 30% by 2010 (Census).

In terms of non-targeted statistics, average lifetime earnings increase over time, since
Table II: Results

<table>
<thead>
<tr>
<th>Targets</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3 peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>90-10 wage diff in 1970 = 1.15</td>
<td>1.20</td>
<td>1.20</td>
<td>1.20</td>
<td>1.20</td>
</tr>
<tr>
<td>Avg log college premium = 0.55</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.61</td>
</tr>
<tr>
<td>Avg 4y college cost to lifetime earnings = 0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>College completion rate in 1970 = 0.14</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>College completion rate in 2005 = 0.30</td>
<td>0.33</td>
<td>0.44</td>
<td>0.54</td>
<td>0.61</td>
</tr>
<tr>
<td>College completion rate in 2005, with noise</td>
<td>0.33</td>
<td>0.38</td>
<td>0.39</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Other Statistics

<table>
<thead>
<tr>
<th>Targets</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3 peers</th>
</tr>
</thead>
<tbody>
<tr>
<td>90-10 wage diff in 2005 = 1.65</td>
<td>1.29</td>
<td>1.32</td>
<td>1.32</td>
<td>1.32</td>
</tr>
</tbody>
</table>

Educational attainment increases. The 90-10 log wage differential also increases, though not as much as in the data.

**Speed of learning:** Figure 6 plots the fraction of the population that is educated in each generation for different numbers of peers.

![Figure 6: Simulation. Investment in education as a function of the number of peers observed.](image)

It is assumed that nobody is educated in generation 0. Under full information, the fraction would be 98% starting with generation 1. If each individual only learns from his or her own experience, learning is quite slow. The addition of each additional signal speeds up
learning substantially. On the other hand, if peers’ wages are observed with noise, learning slows, as shown in the second panel of the figure.

Figure 7 shows that education rises faster in populations with a larger variance of ability, at least in the early generations.

Inequality: Figure 8 shows the evolution of wage inequality over time, as measured by the Gini coefficient (panel a) and by the 90-10 log wage differential (panel b). Inequality first rises, but then once educational attainment reaches at least 30%, it begins to fall. Once educational attainment reaches 100%, inequality returns to the initial level (in period 0, when nobody is educated), since education benefits individuals of all abilities symmetrically.
Mobility: Consider first the canonical measure of relative mobility: the intergenerational income elasticity (IGE), defined as

$$IGE_{t,t+1} \equiv \frac{Cov(\log w_{jt}, \log w_{jt+1})}{Var(\log w_{jt})},$$

(19)

obtained by regressing log child income on log parent income.

Next, consider the the correlation between child and parent percentile rank in the income distribution,

$$Rank_{t,t+1} \equiv Corr(P_{jt}, P_{jt+1}),$$

(20)

obtained by regressing the child’s percentile rank in the income distribution of children on the parent’s percentile rank in the income distribution of parents.

Figure 9 shows the evolution of wage mobility over time, as measured by the IGE coefficient (panel a) and by the rank-rank slope (panel b). Mobility initially declines, as those with higher wages are initially more likely to invest. However, as average wages rise and as beliefs about the value of education become more precise, more agents invest in education regardless of their current earnings, thereby increasing mobility. Note that mobility starts to increase in tandem with the decrease in inequality: the two go hand in hand in the current model. Eventually, as educational attainment reaches 100%, the economy reaches full mobility (zero correlations between parents and children).

Figure 9: Simulation. Mobility over time as a function of the number of peers observed.
3.2 Segregation

In order to quantify the effect of segregation on the speed of learning, we consider communities that only differ in the pool of peers to be sampled by each individual. Otherwise, they are identical (the same distribution of abilities, the same returns to education, etc.). In the most diverse community, each individual draws peers from the entire population of the community. In the segregated community, each individual draws peers from a sample of the population centered around his or her own current wage. The more segregated is the community, the narrower the range of potential peers around the individual’s wage level. In the extreme, the individual only samples from peers who have the same wage as him or her.

**Speed of learning:** Figure 10 shows the speed of learning as a function of the degree of segregation for the case of learning from two peers. More segregation (smaller sized groups within the community) slows down learning. Note, in this simulation, the individuals do not take into account the fact that they are drawing signals from a restricted sample when updating their beliefs. The second panel of the figure shows that the beliefs of individuals in the more segregated community are less precise.

![Figure 10: Simulation. Speed of investment in education as a function of segregation, for the case of learning from two peers. Segregation is measured by the size of an individual’s potential peer group, where the group is centered on the individual’s wage.](image)

**Inequality:** Segregation reduces the median wage, hence it generates an increase in inequality between the integrated community and the segregated community. Because educational attainment is lower in the segregated communities, inequality is also higher. These results are shown in the two panels of Figure 11.
Mobility: Figure 12 shows the evolution of mobility for the case of learning from two peers, for high versus low degrees of segregation: segregation yields lower mobility throughout.

Figure 11: Simulation. Evolution of median wage and wage inequality.

Figure 12: Simulation. Evolution of mobility.
References