Leveraging big data to help restore the American Dream

None of us got where we are solely by pulling ourselves up by our bootstraps. We got here because somebody—a parent, a teacher, an Ivy League crony or a few nuns—bent down and helped us pick up our boots.

—Thurgood Marshall

The “American Dream” means different things to different people, and there are many different ways to measure whether people have achieved it. Raj Chetty suggests a simple statistic that can be measured empirically: the probability that a child born to parents in the bottom fifth of the income distribution reaches the top fifth of the income distribution as an adult. Since by definition only 20 percent of the population can be in the top fifth of the income distribution, the upper bound on this statistic is 20 percent. That is, if the economic circumstances an individual is born into had no effect on economic mobility, then 20 percent of those born into the bottom fifth would reach the top fifth. In the United States, 7.5 percent of those who start out in the bottom fifth of the income distribution reach the top fifth.\(^1\)

This rate is low compared to other developed countries around the world; for example, the equivalent statistic in the United Kingdom is 9 percent, in Denmark 11.7 percent, and in Canada 13.5 percent.\(^2\) This means that the chances of achieving the “American” Dream are almost twice as high if you grow up in Canada.

Differences in economic opportunity within the United States

Policy discussion has looked at these cross-national-differences in mobility, but Chetty warns there are many
issues associated with making such comparisons, especially due to the many differences between countries. For this reason, the focus of this article and his lecture that it summarizes is on variation in economic mobility within the United States, which, it turns out, is even larger than that across countries.

Chetty describes his recent work with Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, which has documented upward mobility rates for 741 metropolitan and rural areas covering the United States, using anonymous earnings data for 40 million children born in the United States between 1980 and 1993. The study is an example of an important trend in economics according to Chetty, which is the application of “big data” to public policy questions. As Figure 1 shows, they find substantial variation in the probability of rising from the bottom fifth to the top fifth depending on where someone grew up. For example, for a person raised in Milwaukee, Wisconsin, that probability is only 4.5 percent, while in San Jose, California, it is 12.9 percent. Some of the variation is regional; for example, the Southeast tends to have considerably lower rates of mobility than the West Coast. Even within regions, however, nearby areas can have very different mobility rates. There is also often considerable variation within metropolitan areas.

Why does upward mobility differ across areas?

Chetty and colleagues’ data indicate that much of the geographic variation in upward mobility can be attributed to the causal effects of childhood environment. That is, the variation in mobility is explained primarily by differences in neighborhoods, schools, and other aspects of a child’s surroundings rather than by demographic differences between the people living in different locations, or by other differences between locations, such as the types of jobs that are available. Chetty and Hendren looked at 8 million families that moved between locations, and made use of variation in the age of children when that move takes place. The results, as shown in Figure 2, illustrate that as the child’s age increases, the benefit of moving to a location with higher income mobility wanes. So, for example, if those who grew up in an area with high income mobility earn $40,000 on average, and those in an area with low income mobility earn $30,000, then children whose families move from the
lower to the higher mobility area when they are 9 years old will earn an average of $37,000. In other words, they get 70 percent of the gain of living in the better neighborhood since birth. This pattern appears to be linear until age 23, meaning that every extra year in the better neighborhood matters equally, not just the earliest years. These results suggest that moving even a teenager to an improved environment can have a substantial effect on his or her adult earnings.

The existence of a relationship between the age of a move and adult income does not necessarily imply that the childhood environment causes the variation in income mobility. For example, it is possible that families who move when their children are younger differ in characteristics that are related to adult income from those who move when their children are older. The researchers cite the results of their comparisons of sibling effects within families as evidence that this relationship is actually causal. For example, they find that when a family with siblings of varying ages moves to a better area, younger siblings do better in adulthood, and the adult income gap is proportional to their age difference.

**What characteristics are correlated with income mobility?**

While evidence indicates that moving children to high-mobility areas provides substantial benefits in the form of greater adult income, that does not explain why some areas have higher mobility than others. Chetty and colleagues identify five factors that are highly correlated with income mobility: segregation, income inequality, school quality, family stability, and social capital.

**Racial and economic segregation**

Racial and economic segregation are associated with significantly less mobility in Chetty’s study. For example, Milwaukee, a city with a high degree of segregation, has very low income mobility. In contrast, Sacramento, where the proportion of people of color is similar to that of Milwaukee, but where the level of integration is much higher, has one of the highest levels of upward mobility in the United States.

**Income inequality**

The data also indicate that areas with a smaller middle class, that is, people who are between the 25th and 75th percentiles of the national income distribution, have much less upward mobility. Chetty noted that this could indicate a direct link between inequality and social mobility. While there is a large difference of opinion on whether it is appropriate for the government to attempt to reduce inequality by redistributing resources in a more equal manner, most Americans do believe in the ideal of social mobility and equality of opportunity. That is, a child’s chance of success should not depend solely on their parents’ economic status. Thus, even those who would not advocate for a reduction in inequality for its own sake might take up the cause if it would increase the probability that anyone could achieve the American Dream.

It is also notable that differences between areas in the number of people who fall in the very top of the income distribution...
are not highly correlated with differences in mobility. For example, the Bay Area in California is home to some of the richest people in the country, but still has very high rates of social mobility.

**School quality**

Areas with high social mobility also tend to be areas with indicators of higher school quality, including higher spending on public schools, smaller class size, and higher test scores conditional on income. This finding is consistent with the idea that human capital is important in determining one’s level of economic success.

**Family stability**

The single strongest correlation with social mobility is the proportion of families that are headed by a single parent; areas with more single parents have substantially lower levels of social mobility. The researchers note that this effect exists even for children of married parents. That is, a child growing up in a two-parent family but in an area with a high proportion of single parents is less likely to move up in the income distribution than if that child lived in an area with a low proportion of single parents, all else equal. Thus, although there is a direct effect of whether a child’s parents are married, there also appears to be an indirect effect of the type of community in which they live.

**Social capital**

The concept of “social capital” became widespread in the 1990s, and was the subject of Robert Putnam’s book *Bowling Alone: The Collapse and Revival of American Community*. Social capital refers to the benefits provided by social networks and other features of social organization. In areas with high social capital, there are likely other people who will help you when needed. One of the measures of social capital used in Putnam’s book is the number of bowling alleys in an area. And in fact, Chetty and colleagues find that the number of bowling alleys is indeed positively associated with rates of social mobility. (Chetty also notes that this finding highlights the distinction between correlation and causality; he does not conclude that building more bowling alleys will increase upward mobility.)

**Policy changes that can improve social mobility**

If more bowling alleys will not boost social mobility in the United States, what will? To find out, the next step is to explore the causal mechanisms behind the correlations, and identify some promising policy changes. Chetty focused on two types of policies: reducing segregation through affordable housing policies; and increasing teacher effectiveness in order to improve school quality. Chetty noted that he was not implying that other factors were not important, or even that these were the two most important factors in increasing social mobility, just that they were a pragmatic choice, given that income inequality, family stability, and social capital have historically proven difficult to change through policy.

**Reducing segregation**

One way to try to reduce racial and economic segregation is to give families housing vouchers and encourage them to move to better neighborhoods. The U.S. Department of Housing and Urban Development’s Moving to Opportunity (MTO) demonstration sought to evaluate the effects of such vouchers. This demonstration, implemented in the mid-1990s, involved 4,600 families with children living in public housing in high-poverty neighborhoods in Baltimore, Boston, Chicago, Los Angeles, and New York. Families who were randomly assigned to the experimental group received housing vouchers that could be used only in areas with poverty rates below 10 percent, and help from a housing-mobility counselor in finding eligible housing. A control group remained eligible for their current housing assistance, but received no additional help through the program. Over a 10- to 15-year follow-up period, the MTO experiment was found to have no significant effects on economic outcomes for parents, and no systematic effects on academic achievement for children. However, Chetty, Hendren, and Lawrence Katz extended this previous work using data on adult economic outcomes for those who had been children at the time of the demonstration, and found that children who moved to a lower poverty area at a young age had substantially better economic outcomes as adults than those who didn’t move. For example, children who moved before the age of 13 from the Martin Luther King Towers public housing development in Harlem, to the lower-poverty area of Wakefield in the Bronx, had earnings that were on average 30 percent higher than children whose families remained in Harlem. The children who moved to a better neighborhood at a young age were also 27 percent more likely to attend college, 30 percent less likely to become single parents, and tended to live in better neighborhoods themselves as adults than the control group. As a result, the movers’ own children are now growing up in better environments—so the effects of the initial intervention appear likely to persist to the grandchildren of the original MTO participants. Chetty, Hendren, and Katz also found, consistent with earlier work, that moving had little effect on children who were older at the time of the move.

These findings suggest that moving to a mixed-income neighborhood improved outcomes for low-income children, but this finding alone is not enough to recommend policy, since it tells us nothing about the overall effects of such a policy on a large scale, including the effects on those who already lived in the more desirable neighborhood. While these types of effects are very hard to detect experimentally, results of the analysis of the effects of growing up in different U.S. counties, described above, show that mixed-income areas result not only in better outcomes for children from low-income families, but also in slightly better outcomes for children from high-income families. This suggests that making a neighborhood more economically integrated is
not a zero-sum proposition, but could potentially increase overall well-being, rather than improving outcomes for some and worsening them for others.

If this is true, then a potential policy implication could be to change the current system of subsidized housing vouchers in a way that encouraged families with young children to move to better neighborhoods. Currently, families seeking a housing voucher often get put on a long waiting list, so they may not have the opportunity to use it to move to a better neighborhood until their children are older. However, there are clearly limits both to the scalability and practicality of such a policy; moving people around cannot be the only solution, though it might be a useful short-term approach. In the long run, policymakers and urban planners could think about how to improve existing neighborhoods, and how to design cities that are more integrated.

Improving school quality

Another strategy that Chetty suggests could increase social mobility is through education policy, specifically increasing teacher quality. Earlier research by Chetty and colleagues found that high-quality teachers can have significant effects on the later earnings and on upward mobility of their students.8 One currently prominent way of measuring teacher quality is to use teacher value-added measures, which gauge how much a given teacher raises students’ test scores on average. To assess the effects of having a high value-added teacher, the researchers used test data from 2.5 million children over a 20-year period (another example of big data research), linked to federal income tax returns in order to obtain earnings and other adult outcomes. They find that replacing (or improving through training) a teacher who is in the bottom 5 percent of the distribution of value added with a teacher of average quality, would increase the undiscounted lifetime earnings of the typical child by $50,000, or $1.4 million for each average-sized class of 28 students.

Upward mobility and economic growth

The last set of results discussed by Chetty suggests that improving opportunities for upward mobility might be desirable based not solely on principles of justice, but also from a perspective of economic growth. In particular, he suggests that one child’s success need not necessarily come at another’s expense. To illustrate this, he discussed one specific pathway to upward mobility: innovation.

Innovation

Using patents for inventions as a proxy for innovation, this analysis uses data on 750,000 patent holders in the United States, linked to tax data. About 2.2 of every 10,000 children born to parents with below-median income will have a patent by the time they are 35 years old.9 In contrast, those who are born to parents in the top 1 percent of the income distribution are 10 times as likely to have a patent by the same age. This gap in rates of innovation related to parental income could be about genetics and the persistence of ability across generations, or it could be related to differences in childhood development.

Figure 3. Patent rates by third grade test scores, for children with low- and high-income parents.

environment as suggested above. To try to identify the cause of this gap, Chetty and colleagues used the same school test data used in the value-added analysis to approximate ability at early ages. This analysis shows that the probability of innovation as an adult is very low for third-graders who are below the 85th percentile of test scores; above that level, the rate of innovation rises sharply. Figure 3 shows this relationship with separate lines for children whose parents have income below the median and those with income above the median. Each data point corresponds to 10 percent of the test score distribution. The two series look very similar right up to the last pair of data points; the children with the highest ability, those in the top 10 percent of the test score distribution, are much more likely to become inventors if they are born to high-income rather than low-income parents.

Thus, even when conditioning on a measure of ability, it appears that there is a large gap in innovation by parental income. This suggests that the difference might be due to differences in the types of resources or environments experienced by children, rather than just differences in ability between children in low- and high-income families. The data also show that as children age, these test score gaps increase, and test scores explain more of the innovation gap. That is, children from low-income families are falling behind in terms of achievement relative to children from high-income families, and this increases over time. Again, this appears to be consistent with the view that the innovation gap could be explained by differences in childhood environments. According to Chetty, this analysis implies that improving equality of opportunity is of interest not only to those who begin at the bottom of the income distribution, but potentially to all families, since increasing the overall amount of innovation could benefit society as a whole.

Policy lessons

Chetty draws three policy lessons from the analyses he described. First, it makes sense to think about issues of social mobility at a local rather than national level. The American Dream appears to be alive and well in some locations, but not in others. Thus, policy should be aimed at increasing social mobility where need is greatest. Second, it is important to focus on improving the childhood environment, and particularly on improving the environment throughout childhood, not just during the earliest ages that receive the most attention in the current policy debate. One short-term solution to improving the childhood environment is to build on the existing subsidized housing voucher programs in order to try to help families move to better areas. A longer-term solution is to improve neighborhoods. Chetty discussed doing this through improving the quality of schools, but noted that there are likely a number of different ways to improve specific neighborhoods. Third, large datasets can be very helpful in both evaluating policy in a rigorous way, and in measuring local progress and performance. In this context, these databases can be used to identify which neighborhoods are in the greatest need of improvement, and which policies actually seem to work.10

Chetty concluded his talk by reiterating his principal finding: the probability of rising from the bottom fifth to the top fifth of the income distribution in the United States is 7.5 percent, less than most other developed countries. Chetty suggests that this disparity presents both an opportunity and a challenge. The local variation across places that are relatively similar, and the fact that outcomes improve when people move to particular areas, suggests that social mobility can actually be changed through policy: this is the opportunity. The challenge, on the other hand, is twofold: first, for researchers to figure out what is causing these differences in mobility across areas; and second, for policymakers to figure out how to effectively implement policies suggested by the research.

---

3 Chetty, Hendren, Kline, and Saez, “Where is the Land of Opportunity?”
4 As the dataset used begins in 1996, there are no data available for children whose families moved when they were younger than nine.
5 Chetty, Hendren, Kline, and Saez, “Where is the Land of Opportunity?”
10 The data described here may be downloaded at www.equality-of-opportunity.org.