An Analysis of Poverty in Massachusetts Cities and Towns

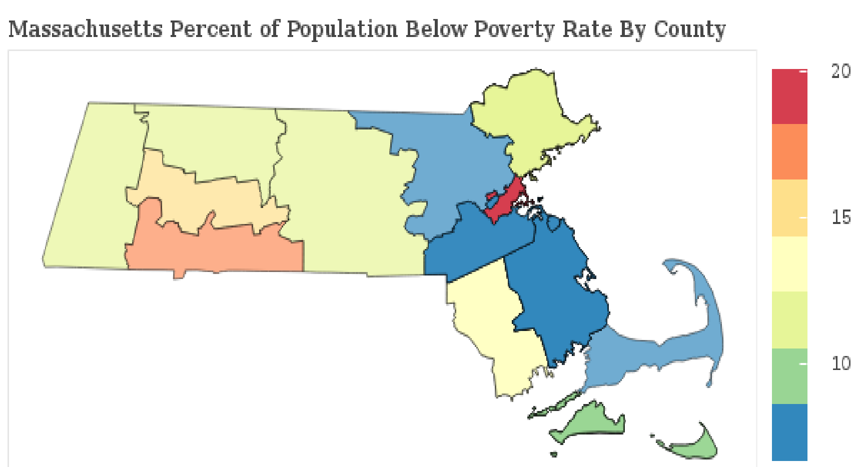
Resource Economics 312
Amanda Carlisle, Christopher Keeler, and Sean Quinn
May 2nd, 2019

I. Introduction

Utilizing data from the U.S. Census Bureau, this study will attempt to build a model to estimate the poverty rate in Massachusetts cities. The poverty rate is formally defined as “the ratio of the number of people (in a given age group) whose income falls below the poverty line; taken as half the median household income of the total population”(1). The interesting part about doing a study like this at the state level is that factors like minimum wage, climate, and economic history are all held constant throughout the state, or have little variation. Because of this, we expect differences in race and education to play a greater role in our model than they would in a national model. This study is of interest because it could potentially be used by state legislature to help people in poverty in Massachusetts, and could provide an explanation as to why some areas are more impoverished than others. It may also highlight and be used to solve problems with income inequality in Massachusetts.

A Snapshot of Current Poverty in Massachusetts

The current poverty rate in Massachusetts is about 11.4%, which ranks 9th lowest in the country. The map below charts the poverty rate in different counties in Massachusetts, and shows that the area with the highest percentage of impoverished people is Suffolk County, where the state’s capital Boston is located (5). While we are analyzing cities and towns, the map of wealth distribution per county provides an overview of poverty rates. Immediately outside of Suffolk County there is the smallest concentration of impoverished people, possibly due to the “white flight” phenomenon of the 1960s, which caused large amounts of wealthy white families to move out of the city and into the nearby neighborhoods. The northernmost parts of the state, as well as western Massachusetts have moderate rates of poverty between 12 and 15 percent.



Literature Review

Almost all literature analyzing poverty agrees that it is a multidimensional issue, which makes it difficult to pin down which independent variables are most important. According to a study done by the World Bank, some of the major determinants of poverty at the household and community level are education, the age structure of household members, the household size, the extent of labor force participation rate, and ethnicity (7). Education has a inverse relationship with poverty, as education allows for more employment opportunities and higher paying jobs, and individuals with more education on average experience greater lifetime earnings (3). Generally, the household size of poorer people is larger than that of more affluent people (7). Labor participation rate also plays an important role, as families with more people in the formal workforce are less impoverished on average. Unemployment rate has also been found to be one of the major determinants in poverty rate. Some studies estimate that increasing the unemployment rate by 1 percent in a given area would result in an increase in the poverty rate by .4 to .7 percent (7).

Ethnicity has a large impact on poverty, as ⅖ of ethnic minorities live in poverty, which is twice the rate of white people (2). The effect is so profound that across all age groups and working statuses, minorities have higher poverty rates than white people of the same category. In *The Colors of Poverty: Why Racial & Ethnic Disparities Persist,* a study done at University of Michigan, it was concluded that the general cause of increased poverty in minorities is "cumulative disadvantage over the life course, as the effects of hardship in one domain spill over into other domains" (Linn and Harris 1). They also found that families in predominantly black or Hispanic neighborhoods have access to less than half the social services that they would have in white neighborhoods, which is why we included total social assistance receipts in the early iterations of our model as an attempt to measure this. This "cumulative disadvantage" will be have to be accounted for in the model, as communities that are predominantly black or Hispanic will likely have higher poverty rates.

Population density also has an interesting impact on poverty rates. According to the census bureau, rural areas record lower overall median incomes, yet they also record lower poverty rates (4). Also, income inequality tends to be lower in rural areas as measured by the Gini coefficient. This is an interesting finding. Looking at the map of poverty in Massachusetts, we can see this to hold true, as the most densely populated areas with the largest cities are the most impoverished.

II. Model Specification

Our model consists of 5 independent variables. We originally had 11, but narrowed it down to the most important causal factors in poverty. Our model can be represented by the function below, where Y is our dependent variable, poverty rate.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + u$$

($\beta_{1...5}$ = *Regression coefficients − population parameters to be estimated.*)

Table 1: Variable names, units, and description

	Variable	Description	Units
X_1	collegegradper	Bachelor Degree or Higher	%
X_2	laborpartrate	Labor Force Participation Rate	%
X_3	foreignper	Foreign Born Persons	%
X_4	popdensity	Population Density	Thousands of people/sq mile
X_5	householdsize	Household Size	average persons/household

Descriptions of Variables

Education (X_1*)*: This variable demonstrates a deep tie to poverty levels due to its impact on people's ability to find and excel in careers that give adequate or living wages. Since 90.3% of people in Massachusetts have a high school diploma (Census Bureau) we measured education based on the percent of the population with bachelor degrees or higher, as we expected a greater amount of variation in the data. We expect this variable to have a negative effect since it is thought that greater education would likely lead to better job prospects leading to greater income and less likely for there people to be in poverty levels.

Labor Force Participation Rate (X_2*)*: We expect this variable to have a strong negative relationship with poverty because those who are unemployed are significantly more likely to have less income then those who are employed, leading to greater poverty. This will be represented as a percent out of the total population above the age of 16.

Foreign Born Persons (X_3): We expect this variable to be positively correlated with poverty. According to our research, areas with relatively high foreign born individuals have increased poverty levels, as on average 18% of immigrants are in poverty and 10% of native-born persons are in poverty (Chapman/ Bernstein). Based on this research, areas with higher percentages of immigrants have decreasing poverty rates, although they are still substantially higher than areas with mostly native-born people . The term "foreign-born", as defined by the Census, is based on individuals who either consider themselves a citizen by naturalization, or do not consider themselves a US citizen and is represented as a percent out of a total population in an area.

Population Density (X_4): We expect this variable to have a positive relationship with poverty levels. According to Karen Tinsley and Matt Bishop "it is well recognized that poverty is unevenly distributed across space (e.g., poverty is concentrated primarily in inner-city America and rural places, while poverty rates are disproportionately low in suburban areas)." In other words, areas with very high or very low population densities are more likely to have greater poverty rates. Since most of the towns and cities included in this study are either urban or suburban areas, it is likely the outcome will have a relatively small positive coefficient. With this being said, there are some rural areas included which may not be appropriately explained due to skewed data towards urban and suburban areas. This variable is measured by population per square mile.

Household Size (X_5): This variable is thought to have a positive relationship with poverty levels and is expressed as average amount of people living in a home within a population. Poorer people tend to have larger households, and larger households mean that the primary earnings are split up among more dependents, which we expect to result in greater poverty (7). Household size data is taken from houses, apartments, and condos and is representative of any person or group of people living together.

Disturbance Term (U): This is the stochastic, or random, disturbance term in our model.

III. Data

The summary statistics for our data are included in **Table 2.** We have also included poverty rate, the dependent variable we are trying to estimate, along with our 5 independent variables. We used a total of 98 observations, made up of a random sampling of towns and cities throughout Massachusetts. Due to the time constraint of the project, we were only able to take a sampling of the 351 towns and cities instead of using the entire population, as we had to manually extract the data from online sources. All data was taken from the US Census Bureau, and was gathered from 2010 through 2018.

Distribution of Poverty Rates

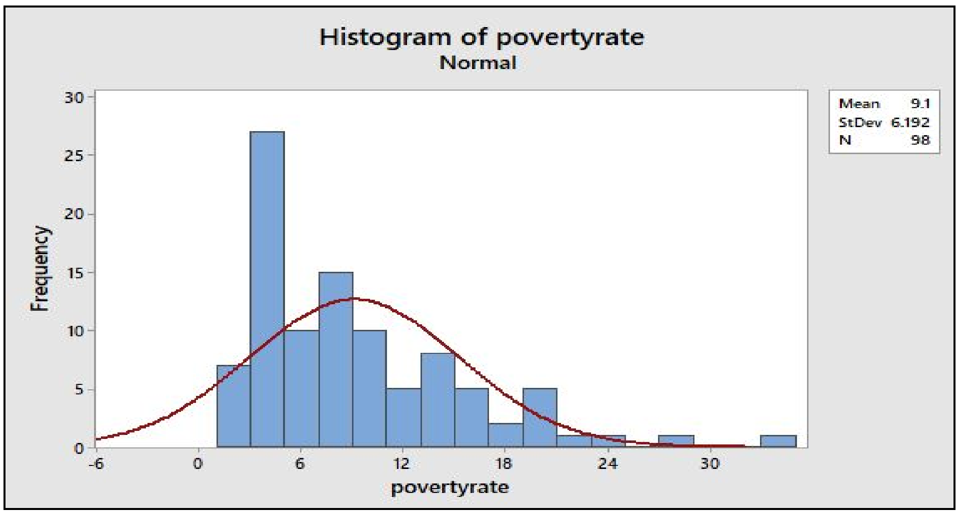


Table 2: Descriptive Statistics (n=98)

Statistics

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
povertyrate	98	0	9.100	0.626	6.192	1.800	4.300	7.550	12.550	33.200
collegegradper	98	0	43.98	1.95	19.35	11.30	28.10	42.95	58.08	83.00
laborpartrate	98	0	67.005	0.531	5.261	48.800	64.550	67.800	70.300	78.900
foreignper	98	0	13.729	0.876	8.671	2.200	7.875	11.400	17.500	45.600
popdensity	98	0	3.092	0.346	3.428	0.010	1.186	1.914	3.432	18.405
householdsize	98	0	2.5347	0.0239	0.2370	2.1200	2.3700	2.5150	2.7200	3.1400

The poverty rates throughout Massachusetts cities and towns vary greatly, with a minimum of 1.8% in Westwood, Massachusetts and a maximum of 33.2% in Amherst, Massachusetts. The distribution of poverty rates, as pictured in the histogram chart, is skewed to the right and has a median of 7.55%. This poverty rate is lower than what we expected to find, as literature estimated the poverty rates to be 11.4% in Massachusetts (5). Some other things to note from the descriptive statistics are the wide variation in percentage of college graduates, which has a minimum of 11.3% and a maximum of 83%, an incredibly large range. Other variables, such as household size, exhibited far smaller variation, as the increase of an entire person per average household in a town or city is an immense difference.

IV. Results

Table 3: Regression results for Poverty Rate

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	1960.41	392.08	20.51	0.000
collegegradper	1	832.12	832.12	43.53	0.000
laborpartrate	1	353.53	353.53	18.49	0.000
foreignper	1	41.03	41.03	2.15	0.146
popdensity	1	144.04	144.04	7.53	0.007
householdsize	1	28.71	28.71	1.50	0.224
Error	92	1758.85	19.12		
Total	97	3719.26			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
4.37241	52.71%	50.14%	45.42%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	45.39	6.91	6.57	0.000	
collegegradper	-0.1551	0.0235	-6.60	0.000	1.05
laborpartrate	-0.3911	0.0909	-4.30	0.000	1.16
foreignper	0.1215	0.0830	1.46	0.146	2.63
popdensity	0.573	0.209	2.74	0.007	2.60
householdsize	-2.64	2.16	-1.23	0.224	1.33

Regression Equation

povertyrate = 45.39 - 0.1551 collegegradper - 0.3911 laborpartrate + 0.1215 foreignper + 0.573 popdensity - 2.64 householdsize

Our final model has an R squared value of 52.71%, which means that 52.71% of the variation of poverty rate is explained by the model. We ran an F test of overall significance, and found that we could reject the null hypothesis, as the critical value is 2.16 and our F value for our model is 20.51. This means that our model is statistically significant at the 1% level of significance significance. 3 of our 5 variables were statistically significant at the 1% level of significance: bachelor degree or higher percentage, labor participation rate, and population density. All of our variables have the effects on poverty rate that we originally expected, except for household size. The final model has no signs of multicollinearity, as all of the variance inflation factors are quite low. Multicollinearity was one of our biggest problems with our earlier iterations, and we will describe how we remedied this problem in the next section. Since it is not time series data, there cannot be autocorrelation. This leaves us with testing for heteroskedasticity, which was a challenge due to the software we used (Minitab) not allowing us to conduct a White's Test. However, we were able to use Excel to perform this test, as shown in **Table 4**. The significance value of this test is .357, well above .05, which means we fail to reject the null hypothesis of homoscedasticity. This means that our model is does not display signs of heteroskedasticity.

Table 4: White's Test Results

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regressio	2	8253.172	4126.586	1.041403	0.356953169
Residual	95	376440.1	3962.527		
Total	97	384693.3			

Parameter Analysis:

College Graduation Percent: The coefficient of this variable is -.1551. The interpretation of this coefficient is: a one percent increase in the percentage of people with a bachelor degree causes a .1551 percent decrease in poverty rate, holding all other variables constant. This is consistent with our original hypothesis and is consistent with economic theory, as an increase in education leads to lower poverty rates.

Labor Participation Rate: The coefficient of this variable is -.3911. The interpretation of this coefficient is: A one percent increase in the labor force participation rate causes a .3911 percent decrease in poverty rate, holding all other variables constant. This is consistent with our original hypothesis and is consistent with economic theory, as an increase in the percentage of the labor force that is employed leads to a decrease in poverty rates.

Foreign Born Persons: The coefficient of this variable is .1215. The interpretation of this coefficient is: A one percent increase in the percentage of foreign born persons causes a .1215 percent increase in poverty rate, holding all other variables constant. This is consistent with our original hypothesis and is consistent with economic theory, as an increase in the percentage of the foreign born persons leads to a increase in poverty rates, in part due to the cumulative disadvantages experienced throughout life by immigrants in America. However, this variable has a larger effect on poverty rates than we originally thought it would have.

Population Density: The coefficient of this variable is .573. The interpretation of this coefficient is: A one thousand person increase in the population density per square mile causes a .573 percent increase in poverty rate, holding all other variables constant. This gives great insight into differences between suburban areas, which may just hold 1000-2000 people per square mile, and urban areas which may hold 5000-10000 people per square mile, and how poverty differentiates between the two. Reasoning for this may have to do with access to affordable housing, affordable food, and other necessities which will more than likely decrease in urban areas where income levels are extremely different between people. This leads to those with less income having to buy relatively expensive food and shelter which may bring about further poverty.

Household Size: The coefficient of this variable is -2.64. The interpretation of this variable is: A one person increase in the number of people per household causes a 2.64 percent decrease in poverty rate, holding all other variables constant. This is the one variable that did not have the effect we expected. We expected that greater household sizes would cause greater levels of poverty, as the income of the primary earners is split up among more dependents. However, the model shows the inverse is true, which might mean that as people become more wealthy in Massachusetts, they are more willing to have more children. One of the studies we used to derive this economic theory was written about global poverty, as people in third world countries have less access to birth control, and often have to have more children to take care them in old age (7). Since Massachusetts is a highly developed area, these factors are likely not very relevant to poverty rates in Massachusetts.

V. Changes to our Model over time

Originally, our model contained up to 11 independent variables. One of the problems we faced was was multicollinearity. We had all the tell tale signs: high variance inflation factors, a high r squared of 79%, high correlation coefficients, and low t values (see **Table 6**). We had high VIFs with percentage of college graduates, percentage of non english speaking, and foreign born persons. We dropped non english speaking, as based on theory it was clear that this would be multicollinear with foreign born persons, as people born outside of America are much more likely to speak another language (2). Looking at **Table 5**, a Pearson coefficient table, these two variables had an incredibly high correlation coefficient of .963, so it made sense to get rid of one of them.

We also realized throughout our process that a number of our variables were not factors that caused poverty, they simply described it. These variables included: median rent, median home value, and median income. We ended up dropping all of these variables from our model, which had a profound impact on our model. We had originally expected median income, median rent, and median home value to have a negative effect on poverty, and while there are strong negative correlations, they are not causal. Lastly, we ended up dropping total social assistance, which was measured in millions of social assistance receipts. We expected this to have a negative effect on poverty rate, as social assistance is supposed to lighten the financial load for impoverished people. Instead, we got a near 0 coefficient, which led us to believe that the effect is not as profound as we originally thought. Like some of our other variables, it was not a causal factor, and more so described poverty in a given area. Looking back this makes sense, as more social assistance to areas with higher poverty rates. Another problem with this variable was that a close to a third of our data points had 0 total social assistance recorded on the Census Bureau, as most impoverished people who need social assistance can't afford to live in many Massachusetts cities and towns altogether.

In an effort to lower our p values and account for more variation in the population, we doubled our sample size in between our second and final installment, from 44 towns and cities to 98 in the final presentation. This, combined with dropping non english speaking percentage, median income, median rent, median home value, and total social assistance, ended up giving us a statistically significant model with no signs of multicollinearity, autocorrelation, or heteroskedasticity.

Table 5: Pearson's Correlation Table

Correlations

	collegegradper	laborpartrate	medhomevalue	medrent
laborpartrate	0.063			
	0.682			
medhomevalue	0.904	0.172		
	0.000	0.265		
medrent	0.527	0.225	0.466	
	0.000	0.142	0.001	
foreignper	0.027	0.360	0.183	0.200
	0.862	0.016	0.234	0.193
totalsocialassis	0.028	0.113	0.080	0.058
	0.856	0.464	0.605	0.706
nonenglishspeak	-0.174	0.306	-0.005	0.087
	0.258	0.043	0.975	0.576
popdensity	-0.035	0.348	0.139	0.156
	0.824	0.020	0.368	0.313
householdsize	0.228	0.246	0.302	0.219
	0.136	0.107	0.046	0.153
medincome	0.775	0.173	0.726	0.311
	0.000	0.262	0.000	0.040

	foreignper	totalsocialassis	nonenglishspeak	popdensity
totalsocialassis	0.288			
	0.058			
nonenglishspeak	0.963	0.255		
	0.000	0.094		
popdensity	0.809	0.458	0.790	
	0.000	0.002	0.000	
householdsize	0.285	-0.138	0.281	0.018
	0.061	0.373	0.065	0.909
medincome	-0.092	-0.126	-0.227	-0.212
	0.553	0.415	0.138	0.167

	householdsize
medincome	0.544
	0.000

Cell Contents
Pearson correlation
P-Value

Table 6: Old Regression Results

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	32.1	11.7	2.75	0.010	
collegegradper	0.2805	0.0895	3.13	0.004	11.72
laborpartrate	-0.428	0.151	-2.84	0.008	1.47
medhomevalue	-0.000033	0.000009	-3.85	0.001	7.23
medrent	-0.00027	0.00161	-0.17	0.867	1.70
foreignper	-0.162	0.331	-0.49	0.627	34.45
totalsocialassist	0.000000	0.000000	1.44	0.159	1.34
nonenglishspeak	0.263	0.229	1.15	0.260	36.25
popdensity	0.000403	0.000285	1.41	0.168	4.39
householdsize	4.95	3.54	1.40	0.172	2.79
medincome	-0.000122	0.000037	-3.26	0.003	5.82

Regression Equation

povertyrate = 32.1 + 0.2805 collegegradper - 0.428 laborpartrate - 0.000033 medhomevalue - 0.00027 medrent - 0.162 foreignper + 0.000000 totalsocialassist + 0.263 nonenglishspeak + 0.000403 popdensity + 4.95 householdsize - 0.000122 medincome

VI. Summary and Conclusion

Poverty is a multifaceted issue, with a plethora of variables contributing to it. In this study five different variables were identified: college graduation percentage, labor force participation rate, foreign persons, population density, and household size. While these variables did not capture all of variation in poverty, as there are likely numerous other factors that contribute, we found that our model overall was statistically significant at the 1% level, with 3 of the 5 variables also being significant: bachelor degree or higher percentage, labor force participation rate, and population density. Each one can be explained through theory when discussing their roles in society and their economic impact on a communities welfare. The data analyzed in this study provides an understanding in why graduation percent and labor force participation rates are tied with a negative relationship on poverty rate which consisted of coefficients -.1551, -.3911, while population density would have a positive relationship on poverty rates with a coefficient of .573. These are not surprising figures when considering our predicted variable relationships earlier in this study.

When looking at the other two variables of foreign persons and household size which were used in this study, it was disappointing to see that neither carried significance at the 1% level. Although this may have been disappointing, neither of these variables should be ruled out when discussing poverty rates, as they are economic theory does show them to be important when considering poverty rates. Since our model was at the state level, a national study might show factors such as climate, opportunity programs for immigrants, and minimum wage, which at the state level are essentially kept constant, lead to the variables of foreign persons and household size to have an increase in significance on poverty rates. On the other hand, if this study had the time to gather more data points on Massachusetts, we would have likely seen the p values for these variables go down, possibly making them statistically significant.

By honing in on the variables found to be significant, state legislature can implement policy that leads to immense benefits to those in need and currently at the poverty level. For example, one of the most important factors was having a bachelor's degree, so in order to help people in poverty, the state legislature could start by increasing access to college. By increasing the accessibility and affordability of a college degree, the state could take one step forward in decreasing poverty in the state of Massachusetts. They can do this by offering more need-based grants and decreasing in-state tuition rates. The state legislature should also focus on college graduation and retention, not simply admittance to college, as our study showed that specifically the completion of bachelor degree program lead to lower poverty rates. As college degrees become increasingly expensive, they are simultaneously becoming increasingly important. A students ability to pay off their student loans is dependant on their ability to graduate. The state government could offer some sort of first-year transition programs to provide academic, financial, career advising to those enrolled in college in order to increase graduation rates.

We also found labor force participation rates to be important in combating poverty, as the higher the labor force participation rate is in a community, the lower poverty rates are. Technological advancements in today's economy has led to a shift in employer's demand for specific technical skills. This shift has left many workers left unprepared, discouraged, and unemployed by the mismatch in skills desired. For those who face this problem, the government should provide career guidance programs and promote skill acquisition in high demand jobs in today's economy. By providing job specific training, the government could ensure that no one is left discouraged and everyone has an equal opportunity to find employment in today's competitive workforce, thus improving the labor force participation rate. Programs could be developed in elementary, middle, and high schools to promote lifelong learning in young workers. If the government were to stress the importance of staying up-to-date with the skills desired in today's workforce, many people will not be left in fear of job termination. The state legislature could also increase labor force participation rate by increasing spending on public works projects that would create jobs and increase employment levels. Lastly, population density proved to be an important factor, showing that places with higher population densities had much higher poverty rates. This could be combated by implementing the systems talked about above, but focusing them specifically in areas with higher population densities.

VII. Resources

1. OECD (2015), *In It Together: Why Less Inequality Benefits All*, OECD Publishing, Paris,https://doi.org/10.1787/9789264235120-en.
2. Kenway, Peter, and Guy Palmer. "Poverty Among Ethnic Groups How and Why Does It Differ." *Joseph Rowntree Foundation* , New Policy Institute, Mar. 2007, www.poverty.org.uk/wp-content/uploads/2018/09/ethnicity.pdf.
3. Majumder, Sacchidanand & Biswas, Soma. (2017). The Role of Education in Poverty Alleviation: Evidence from Bangladesh. 8.
4. Nasser, Haya El. "Poverty Rate Lower in Rural America Than Urban Centers." *Census Bureau QuickFacts*, United States Census Bureau, 18 Sept. 2018, www.census.gov/library/stories/2017/11/income-poverty-rural-america.html.
5. "Poverty in Massachusetts." *Apply For Welfare*, www.welfareinfo.org/poverty-rate/massachusetts/.
6. Lin, Ann Chih, and David R Harris. "Policy Brief." *National Poverty Center*, University of Michigan, Jan. 2009, www.npc.umich.edu/publications/policy_briefs/brief16/PolicyBrief16.pdf.
7. "Understanding the Determinants of Poverty." *Understanding the Determinants of Poverty*, Poverty Manual, 8 Aug. 2005, siteresources.worldbank.org/PGLP/Resources/PMch8.pdf.
8. Chapman, Jeff, and Jared Bernstein. "Immigration and Poverty: How Are They Linked?" *U.S. Bureau of Labor Statistics*, Monthly Labor Review, Apr. 2003, www.bls.gov/opub/mlr/2003/04/art2full.pdf.
9. Tinsley, Karen, and Matt Bishop. "Poverty and Population Density: Implications for Economic Development Policy." *Journal of Higher Education Outreach and Engagement* , vol. 11, no. 1, 2006, pp. 195–208., doi:10.3897/bdj.4.e7720.figure2f.