Spending Patterns and Cost of Living for Younger versus Older Households

Payment Card Fraud Rates in the United States

Why Aren't More People Working in Low- and Moderate-Income Areas?
Spending Patterns and Cost of Living for Younger versus Older Households
By Jun Nie and Akshat S. Gautam

Payment Card Fraud Rates in the United States
By Fumiko Hayashi

Why Aren’t More People Working in Low- and Moderate-Income Areas?
By Kelly D. Edmiston
Spending Patterns and Cost of Living for Younger versus Older Households

By Jun Nie and Akshat S. Gautam

Aggregate measures of inflation can mask large differences in the actual cost of living faced by households with different spending patterns. For example, older houses typically spend more on health-related services, while younger households spend more on education. If prices in the health-care and medical services sectors rise at a faster pace than prices in the education sector, older households may, in turn, experience a higher inflation rate than younger households.

Jun Nie and Akshat S. Gautam use a rich household-level expenditure data set along with price data to measure and examine differences in spending patterns and the cost of living across different age groups. They find that older households in general have faced slightly higher inflation rates than younger households over the past four decades due to health-related expenses. However, they also find that the inflation gap between older and younger households has narrowed significantly over the same period as the inflation rate of health-related expenses has declined.

Payment Card Fraud Rates in the United States

By Fumiko Hayashi

The United States has lagged somewhat behind other countries in implementing steps to mitigate payment card fraud, such as chip card technology and personal identification numbers. Small delays in implementing fraud mitigation strategies could translate to large fraud losses relative to other countries. Although comparing fraud rates across countries can be challenging, Fumiko Hayashi examines payment card fraud rates in the United States along with three countries with the best available data—Australia, France, and the United Kingdom—and finds that the United States has the highest overall fraud rate. Even after migrating to chip card technology, the United States has a significantly higher in-person fraud rate than all three countries but a lower remote fraud rate than Australia and France. Fewer safeguards and differences in prevalent types of transactions may help explain this.
Why Aren’t More People Working in Low- and Moderate-Income Areas?

By Kelly D. Edmiston

Despite overall strength in the U.S. labor market, employment in low- and moderate-income (LMI) communities lags behind non-LMI communities. This employment gap is persistent and has increased over time; as of 2017, 35 percent of residents in LMI communities age 18–64 were not working compared with 24.9 percent in non-LMI communities.

Kelly Edmiston uses a formal text analysis of a unique set of survey comments to examine prominent “employment barriers” in LMI and non-LMI communities. He finds that lower educational attainment and lack of access to transportation and childcare are among the most prominent barriers to employment and are especially prevalent in LMI communities. Although public assistance, disabilities, and chronic health conditions are considerably more prevalent in LMI communities, they are not especially prominent barriers in the text analysis.
Economists often use measures of inflation—the percent change in the aggregate price level in a given period—to estimate changes in the cost of living. For example, an annual inflation rate of 2 percent means that the average household will spend 2 percent more to purchase the same basket of goods this year than in the previous year. However, this aggregate measure can mask large differences in the actual cost of living faced by households with different spending patterns. Older households, for example, typically spend more on health-related services, while younger households spend more on education. If prices in the health-care and medical services sector rise at a faster rate than prices in the education sector, older households may, in turn, experience a higher inflation rate than younger households.

Measuring possible differences in the cost of living across age groups requires a comprehensive picture of these groups’ spending across expenditure categories as well as how prices in these categories change over time. We use the Consumer Expenditure Survey, the most comprehensive household-level expenditure data set in the United States, to measure the spending patterns of households at different ages. After exploring these differences across age groups, we then combine the expenditure data with price data from the Bureau of Labor Statistics to examine differences in the cost of living faced by different age groups.
Our results suggest that older households in general have faced slightly higher inflation rates than younger households over the past four decades. This is mainly because older households spend relatively more on health-related expenses, which have had a higher inflation rate than expenses such as transportation and leisure, on which younger households spend relatively more. In addition, we find that the inflation gap between older and younger households has narrowed significantly over the last four decades as the inflation rate of health-related expenses has declined. The difference in spending patterns of older and younger households has remained relatively stable over time and contributed little to the declining inflation gap.

Section I discusses related research and the data used in the analysis. Section II highlights that older households spend more on health, rent, and household goods and services, while younger households spend more on education, communication, transportation, and leisure. Section III reports the implied inflation gap between younger and older households and demonstrates that this gap has narrowed over time.

I. Related Literature and Data

Total household spending accounts for nearly 70 percent of U.S. GDP, suggesting changes in the spending patterns of households or the age composition of the U.S. population may have macroeconomic implications. Researchers therefore have used various data sets to explore the dynamics of household consumption across age profiles. In general, consumer spending is “hump-shaped” over the life cycle: spending ramps up in early adulthood, peaks around age 40 to 50 and then declines with age (Attanasio and Weber 1995; Gourinchas and Parker 2002; Villaverde and Kruger 2007). This hump-shaped spending pattern may just reflect that earnings and wealth are also hump-shaped over the life cycle, as changes in consumption usually follow changes in income and wealth (Wolff 1992; Huggett 1996). However, declining expenditures in old age may also reflect reductions in work-related expenses and spending on items such as food away from home, which tend to decrease as people age and retire (Aguila, Attanasio, and Meghir 2011; Hurd and Rohwedder 2008). Indeed, Aguiar and Hurst (2013) disaggregate nondurable expenditures into more detailed consumption categories and find that the decline in spending on nondurable goods after middle age is essentially
driven by three categories: food, nondurable transportation, and clothing/personal care.¹

To provide a more complete picture of both the composition and patterns of household spending across different age groups, we use the Consumer Expenditures Survey (CEX) data set from the Bureau of Labor Statistics (BLS). The CEX contains the most detailed information on household spending in the United States and is used extensively by researchers and policymakers alike. However, constructing a consistent household-level panel data set based on the CEX is challenging. Unlike other traditional macroeconomic data sets, CEX data are released in different data files, and their formats and structures vary across different years. In addition, the CEX has undergone numerous changes to its file structure and survey design over the years, requiring researchers to have a clear understanding of where each variable is stored and how to merge data files with different formats. Furthermore, most household data in the CEX files are stored at a highly disaggregated level—specifically, at the Universal Classification Code (UCC) level.² These UCCs are used to construct aggregate spending and income categories, but they often change year to year due to the deletion of old UCCs or the addition of new ones.

To address these challenges, we examine the UCCs across different years and construct expenditure categories that are consistent in their definition. Defining categories in a consistent way allows us to construct a data set that covers 36 years from 1983 to 2018 and contains expenditure information for roughly 7,000 households each year.³

This data set, in turn, allows us to make new contributions to a wide body of research on household spending patterns. For example, although Attanasio and Weber (1995), Gourinchas and Parker (2002), and Villaverde and Kruger (2007) study differences in spending at the aggregate level, we examine more detailed consumption categories to assess how they contribute to the differences at the aggregate level. In addition, with more than 600 expenditure items in the CEX, we provide a more complete picture on spending patterns than Hurd and Rohwedder (2008), who use the Health and Retirement Survey to focus on a particular age group of households. Finally, we cover a larger set of consumption categories and focus on a longer time horizon than Aguier and Hurst (2013) and Aguila, Attanasio, and Meghir (2011), who also
use disaggregated CEX data. We then combine our disaggregated price data to construct age-specific inflation rates and quantify how these differing consumption patterns may have led to differences in the cost of living for different age groups.

II. Spending Patterns by Age Group

As the CEX data from the BLS contain a wide range of spending categories—including some nonconsumption categories, such as spending on mortgages and insurance premiums—we extend Blundell and others’ (2008) definition to more recent years to isolate household consumption expenditure items. In addition, we also include spending on health care and education, two categories likely to differ across age groups, on our list. We then divide households into three age groups to explore their spending behavior: younger households (those with a household head age 29 or younger), middle-age households (those with a household head age 30–60), and older households (those with a household head age 61 or older).

Chart 1 shows that at the aggregate level, household spending is hump-shaped over the life cycle, consistent with Attanasio and Weber (1995). In particular, the chart shows that middle-age households on average spent around $60,000 (measured in 2012 dollars) per year from 1983 to 2018, about $21,000 more than the average spending for younger households ($38,500) and about $13,000 more than the average spending for older households ($47,200). As mentioned in the previous section, this hump-shaped expenditure pattern may simply reflect the co-movement of consumption with households’ income and wealth, which are also hump-shaped over the life cycle. However, it may also reflect changes in households’ spending preferences as they age.

To account for potential shifts in spending categories over time, we next break down households’ spending into six major categories. Specifically, we follow the BLS in combining our 600 CEX items into six major spending categories: health, household goods and services, rent, education and communication, transportation and leisure, and food. As Table 1 shows, each of these six categories includes multiple subcategories. For example, “transportation and leisure” includes around 300 underlying UCCs, while “household goods and services” includes about 200 UCCs.
**Chart 1**

Average Real Spending across Age Groups, 1983–2018

Sources: BLS, Inter-university Consortium for Political and Social Research (ICPSR), and authors’ calculations.

**Table 1**

Summary of Six Major Categories

<table>
<thead>
<tr>
<th>Major categories</th>
<th>Expenditures included</th>
<th>Approximate UCCs included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health</td>
<td>Health insurance, medical equipment, and medical services</td>
<td>60</td>
</tr>
<tr>
<td>Household goods and services</td>
<td>Household furnishing and operations, utilities, personal care, miscellaneous spending</td>
<td>200</td>
</tr>
<tr>
<td>Rent</td>
<td>Rent (including owner-equivalent rent)</td>
<td>15</td>
</tr>
<tr>
<td>Education and communication</td>
<td>Education and communication services (including telephone services, computers, and electronics)</td>
<td>40</td>
</tr>
<tr>
<td>Transportation and leisure</td>
<td>Vehicles, public transportation, gasoline and fuel, food away from home, apparel, alcoholic beverages, recreation</td>
<td>300</td>
</tr>
<tr>
<td>Food</td>
<td>Food at home</td>
<td>5</td>
</tr>
</tbody>
</table>

Source: BLS and ICPSR.
Chart 2 plots average real spending across these major categories for the 1983–2018 period and shows that spending indeed differs by age across categories. Although middle-age households had the highest spending in nearly all categories, older households spent the most on health. Although younger households spent less than middle-age households in all categories, younger households outspent older households on education and communication and on transportation and leisure.

As total spending differs across age groups, comparing absolute spending levels in each category may be misleading. Therefore, Chart 3 shows the share of spending in each category for all three age groups from 1983 to 2018. For the first three categories (health, household goods and services, and rent), spending shares increase with age and are not hump-shaped. For example, health spending accounts for about 4 percent of total spending among younger households, 6 percent of total spending among middle-age households, and 13 percent of total spending among older households. In contrast, for the next two categories, education and communication and transportation and leisure, the spending shares decrease with age. The spending shares for food seem to be the same across age groups. This simple disaggregation highlights that spending patterns vary across ages and categories.
As a check for validity, we compare the average spending shares constructed from our data set with the corresponding shares of similarly defined categories published by the BLS (in the Consumer Expenditure tables) for the 2015–18 period, the most recent years for which data are available. As Chart 4 shows, the shares from the two data sets are comparable across categories.

One advantage of our data set over the BLS set is that we can aggregate the underlying spending categories up to different levels. To compare the spending shares of older and younger households in more detail, we aggregate the UCC-level data up to the 17 subcategories that make up the six categories shown in Table 1.6

Chart 5 shows the average difference in spending shares between older and younger households for all 17 categories across our sample years (1983–2018), where a positive difference indicates that older households spent more than younger households in that category. This difference varies from a level above 4 percentage points to below –8 percentage points, suggesting again that younger and older households’ spending differs considerably across expenditure categories. For example, older households’ share of spending on health insurance is about 4.9 percentage points higher than the share for younger households. In contrast,
Chart 4
Comparison of Our Shares versus BLS Shares
(Average, 2015–18)

Sources: BLS, ICPSR, and authors’ calculations.

Chart 5
Average Difference in Spending between Older and Younger Households for 17 Categories, 1983–2018

Sources: BLS, ICPSR, and authors’ calculations.
older households’ share of spending on vehicle-related expenses is about 8.5 percentage points lower than the share for younger households.

Although the average difference shows clear differences in spending shares between older and younger households, it does not reveal how these shares may have changed over time. To answer this question, we examine a time series of our data. The upper panels of Chart 6 show that spending shares on health, household goods and services, and rent were higher for older households (orange lines) than younger households (blue lines) for all years in our sample. Although the share of spending on health and rent has risen over time for all three age groups, the share of spending on household goods and services has declined. The bottom panels of Chart 6 show that spending shares on food, education and communication, and transportation and leisure were higher for younger households than older households for most years in our sample. Although the share of spending on education and communication has risen over time for all age groups, the share of spending on transportation and leisure has declined. Finally, the share of spending on food has remained relatively stable over time for all age groups, though the share for older households has declined slightly.

Overall, decomposing aggregate expenditures into major components uncovers large differences in spending patterns across age groups. These differences are clear in both absolute levels and in relative shares. In addition, spending shares for major categories show common trends across age groups, leaving differences in the spending shares across age groups relatively stable over time.

III. Implied Inflation Rates for Different Age Groups

The large differences in expenditures across age groups could translate to different inflation rates faced by households in these age groups. For example, if prices increase more quickly for goods and services that primarily older households consume, the inflation rate may be higher for older households than for younger households. To assess this possibility, we follow Hobijn and Lagakos (2005) and McGranahan (2006) and combine relevant subcategories’ CPI price data from the BLS with the expenditure data constructed in the previous section to measure age-specific inflation rates. Specifically, we calculate the inflation rate for a particular age group at time \( t \) as follows:
Chart 6
Trends in Spending Shares of Six Major Categories over Time across Age Groups, 1986–2018

Note: Panels begin in 1986 because we plot the three-year moving averages to show smoothed trends.
Sources: BLS, ICPSR, and authors’ calculations.
where \( age \) refers to the particular age group, \( N \) represents the number of consumption categories (six in our case), \( s_{i,t}^{age} \) represents the average share of spending on consumption category \( i \) relative to total spending for households in age group \( age \), and \( \pi_{i,t} \) denotes the inflation rate of consumption category \( i \) in year \( t \).

Consistent with prior research, we assume that different households face the same price for the same expenditure item even though in reality, people may purchase the same item at different prices. For example, more patient households may be able to purchase the same car at a lower price than households who have less time to shop around. We make this assumption mainly because we lack the data to measure differences in prices.

Chart 7 shows that the average inflation rate is higher for older households than for younger households, though the difference is not large. Specifically, the average inflation rate for younger, middle-age, and older households is 2.46, 2.54, and 2.78 percent, respectively. In other words, older households face a 0.32 (2.78 – 2.46) percentage point higher inflation rate than younger households. Accumulated over a 20-year horizon, the cost of living has increased around 10 percentage points more for older households than younger households.

These differences in cost of living could be the result of changes in the inflation rate of certain categories or changes in each age group’s spending shares on these categories. To illustrate this, we use equation (2) to express the difference in the inflation rate between older and younger households as:

\[
\pi_{i,old}^{t} - \pi_{i,young}^{t} = \sum_{j=1}^{N} \left( (s_{i,j}^{old} - s_{i,j}^{young}) \cdot \pi_{i,t} \right)
\] 

where \( s_{i,j}^{old} - s_{i,j}^{young} \) is the difference in spending shares for subcategory \( i \) between older and younger households and \( \pi_{i,t} \) is the inflation rate for that subcategory in a given year \( t \). This expression shows that the larger the gap in the spending share, the larger the category’s contribution to the inflation gap. In addition, the expression shows that if a difference in spending shares between the two age groups is positive, a higher inflation rate for that category will lead to a larger contribution to the inflation gap from that category.
The results from this decomposition show that the higher inflation rate experienced by older households has been largely driven by their spending in three subcategories: health, rent, and household goods and services. The first bar in Chart 8 shows that the health, rent, and household goods and services categories contributed 0.42, 0.17, and 0.12 percentage points, respectively, to the average inflation gap in the 1984–2018 period, which was partly offset by the transportation and leisure (−0.29 percentage point) and education and communication (−0.08 percentage point) categories. Food did not contribute much to the inflation gap, as the spending shares were about the same for older and younger households. Adding up the contributions from these different components yields a total inflation gap of 0.32, illustrated by the light blue box in the first bar.

The second through fifth bars in Chart 8 show that the inflation gap between older and younger households has shrunk over the last 40 years. In general, the shrinking inflation gap is due to declining contributions from all categories, though the contribution from the health category declined the most over the last four decades (from 0.57 percentage point in 1980 to 0.27 percentage point in 2010). In addition, the general decline in contributions across categories is due to falling inflation rates and not due to a decline in spending differences between

---

**Chart 7**

Average Inflation Rate by Age Group, 1984–2018

Sources: BLS, ICPSR, and authors’ calculations.
older and younger households. As Table 2 shows, the inflation rates for all categories have declined from the 1980s to the 2010s. In particular, the health and education categories saw the largest inflation declines of around 3.9 and 4.4 percentage points, respectively. However, the health category has made a larger contribution to the change in the overall gap because the difference in health-related spending between the two age groups is much larger than the difference in education and communication spending.

Finally, Chart 9 shows that the inflation gap between older and younger households tends to shrink as the overall inflation rate increases. Indeed, the correlation between the inflation gap and overall CPI inflation has been around −0.4 over the last 40 years, though it strengthened to −0.7 from 2000 to 2018. The negative correlation between the inflation gap and the overall inflation rate is mainly due to the fact that younger households spend more on transportation and leisure, a category that tends to see larger price increases than other categories when overall inflation is rising. As younger households spend more on transportation and leisure than older
Table 2
Average Inflation Rates for Six Major Categories by Decade (Percent)

<table>
<thead>
<tr>
<th>Decade</th>
<th>Health</th>
<th>Household goods and services</th>
<th>Rent</th>
<th>Education and communication</th>
<th>Transportation and leisure</th>
<th>Food</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>6.79</td>
<td>2.97</td>
<td>4.95</td>
<td>4.44</td>
<td>2.19</td>
<td>3.84</td>
</tr>
<tr>
<td>1990</td>
<td>5.37</td>
<td>2.72</td>
<td>3.31</td>
<td>2.79</td>
<td>1.84</td>
<td>2.84</td>
</tr>
<tr>
<td>2000</td>
<td>4.25</td>
<td>2.62</td>
<td>3.03</td>
<td>2.10</td>
<td>2.23</td>
<td>2.75</td>
</tr>
<tr>
<td>2010</td>
<td>2.93</td>
<td>0.77</td>
<td>2.46</td>
<td>0.01</td>
<td>1.52</td>
<td>1.22</td>
</tr>
</tbody>
</table>

Sources: BLS, ICPSR, and authors’ calculations.

Chart 9
Inflation Gap between Older and Younger Households and Headline CPI-U, 1984–2018

Sources: BLS, ICPSR, and authors’ calculations.
households, their inflation rate rises faster relative to older households, thereby reducing the inflation gap.

To summarize, by using the expenditure shares to construct age-specific inflation rates, we find that compared to the large differences in spending shares across different age groups, the implied inflation gap is much smaller. In addition, the gap has declined in recent decades, as inflation rates have generally declined for various components.

**Conclusions**

Headline inflation statistics may mask differences in the cost of living faced by different age groups. However, measuring differences in these groups’ relative cost of living requires detailed data on their spending across expenditure categories. We exploit a rich household-level expenditure data set to provide a comprehensive picture of younger, middle-age, and older households’ spending patterns as well as how their inflation rates have changed over time. We find that older households have very different spending patterns than younger households. In particular, we find that older households spend more on health and medical services, household goods and services, and rent, while younger households spend more on education and communication and on transportation and leisure. In addition, we find that the shares of household spending on health, rent, and education and communication have risen for all age groups over the last 40 years, while the shares of spending on household goods and services and on transportation and leisure have declined.

To explore the implications of these different spending patterns on households’ relative cost of living, we combine our expenditure data with subcategories’ price data. We find that older households in general face slightly higher inflation rates than younger households, but the difference has narrowed significantly over the last four decades.
Endnotes

1 Nondurable transportation includes transportation expenses such as gasoline and vehicle repair but excludes spending on durables such new or used cars and trucks.

2 Overall, about 800 UCCs summarize households’ relevant information on spending, income, demographics, assets, and so on. A given UCC thus may uniquely identify expenditures such as groceries, footwear, meals at restaurants, and alcoholic beverages, or income information such as the amount received in transfers, wages and salaries, and financial dividends.

3 The CEX data are available from 1980. However, the first three years lack information on important variables such as owner-equivalent rent. To be consistent, we therefore start our data set in 1983. The official number of households surveyed by the BLS every year has been around 12,000 in recent years. Usable information can be extracted from about 4,000 to 8,000 households every year, with earlier years having fewer households. For the period 1980–95, we get the CEX data from the Inter-university Consortium for Political and Social Research (ICPSR), while the data from 1996–2018 come from the BLS.

4 In robustness analysis, we use alternative age thresholds to define different age groups. Our main results still hold qualitatively.

5 We separate rent from other non-rental household operating expenses included in the “household goods and services” category because rental expenses are less discretionary—that is, households have less control over changing them—than other spending categories. For details on how the BLS aggregates CEX categories, see the CEX-ISTUB hierarchy available at the BLS website. When constructing the expenditure categories, we also need our defined categories to match the relevant price indexes in Section III.

6 Since we later use these underlying categories to construct age-specific inflation rates, we try to match our 17 categories to the BLS’s underlying CPI subcategories.

7 Equation (1) is a weighted average of different subcategories’ inflation. This is an approximation of the growth of the aggregate price. We adopt this formula as it is easier to explain. We construct the price indexes for the six subcategories following a similar formula:

\[ \pi_t^m = \sum_{k=1}^n s_{kt} \cdot \pi_{kt}, \]

where \( n \) denotes the number of subcomponents in category \( m \), \( s_{kt} \) denotes the share of spending in subcomponent \( k \) relative to total spending in \( m \), and \( \pi_{kt} \) is the inflation rate of subcomponent \( k \) in year \( t \). The price information for these subcomponents come from the similar underlying component indexes published by the BLS.
References


Payment Card Fraud Rates in the United States Relative to Other Countries after Migrating to Chip Cards

By Fumiko Hayashi

Although the payment industry around the world has taken major steps to mitigate payment card fraud, the United States has lagged somewhat behind. In the 2000s, many countries adopted or began migrating to a chip card technology called “Europay, Mastercard, and Visa” (EMV) to mitigate fraud from counterfeit cards used for in-person and ATM (or “card-present”) transactions. However, the U.S. payment industry did not begin migrating to EMV technology until 2015. In addition, while other countries require chip card users to input personal identification numbers (PINs) to prevent fraud from lost or stolen cards, the United States has yet to adopt this additional safeguard as standard practice, especially for credit card transactions.

These different fraud mitigation strategies may translate to differences in payment card fraud rates. However, comparing fraud rates across countries is challenging for a few reasons. First, fraud rates after U.S. EMV migration were not available until recently. Second, while some countries report fraud values, they do not report fraud rates; constructing the latter would require detailed transaction data. Third, the level of detail of available fraud rates varies across countries, making it difficult to identify where and why differences in fraud rates occur.

In this article, I compare U.S. payment card fraud rates to fraud rates in three countries with the best available data—Australia, France,
and the United Kingdom—and assess what might explain the differences. I find that even after EMV migration, the United States has a significantly higher in-person fraud rate than all three countries but a lower fraud rate for phone, mail, and internet transactions (remote) than Australia and France. Factors explaining the higher in-person fraud rate include U.S. cardholders’ greater tendency to use credit cards compared with cardholders in other countries, the U.S. payment industry’s late migration to EMV, and EMV implementation without a strong card verification method, such as PINs. The United States’ lower remote fraud rate may be partly explained by a smaller fraction of remote payments made at foreign merchants relative to domestic merchants.

Section I discusses challenges to mitigating fraud in the United States. Section II describes what fraud data are collected in the United States and other countries. Section III shows differences in in-person, remote, and overall fraud rates between the United States and other countries and provides potential factors explaining those differences.

I. Challenges to Mitigating Fraud in the United States

Relative to other developed countries, the United States has historically been slow to implement fraud mitigation measures for both in-person and remote transactions. For example, the United States was one of the last developed countries to migrate to EMV chip technology to mitigate counterfeit fraud in the card-present environment. The United States has fallen behind other countries in adopting stronger authentication technologies to mitigate remote fraud as well. France and the United Kingdom, for example, have progressively adopted authentication technologies, such as 3-D Secure (3DS), since the late 2000s. In the United States, however, card issuers are not expected to start supporting a new version of 3DS called EMV-3D Secure until late 2019. This delay in particular may have implications for the overall fraud rate: in general, the remote fraud rate is significantly higher than the in-person fraud rate, and the share of remote payments has been increasing.

Although all countries need to overcome coordination problems in implementing large-scale fraud mitigation measures, such as EMV chip technology and 3DS, the United States may face greater challenges than other countries. First, the U.S. payment industry is highly complex. More than 10,000 financial institutions issue debit cards, many of which
issue credit cards as well. Millions of merchants, billers, and other businesses accept payment cards. Moreover, many card networks and payment service providers process transactions and offer services to mitigate card fraud. Compared with the United States, other countries have fewer card issuers, merchants, card networks, and service providers.

Second, U.S. public agencies, including the Federal Reserve, lack explicit power to regulate payment systems. Although the Federal Reserve plays an active role in improving security in check, automated clearinghouse, and wire systems, it has little involvement in the payment card system. With debit cards, the Board of Governors of the Federal Reserve System regulates only debit card routing and interchange fees received by large debit card issuers. In contrast, governments or central banks in other countries have regulatory power and thus require or pressure private-sector participants to implement fraud mitigation measures. For example, the Banque de France, whose mandate includes security measures for payment cards, led the nationwide adoption of EMV chip technology and 3DS (Stervinou 2015). The Reserve Bank of Australia (RBA), which has explicit payment regulation power, recently encouraged industry participants to implement a coordinated strategy to mitigate remote fraud (Reserve Bank of Australia 2018). And in the European Union, the revised Payment Services Directive required strong customer authentication for electronic payments starting September 14, 2019.2

Third, participants in the U.S. payment card industry may not have strong incentives to mitigate fraud. U.S. card issuers receive significantly higher revenues from interchange fees charged to merchants relative to fraud losses than other countries, which may make them less sensitive to fraud. One reason for the higher interchange fees in the United States is that the United States regulates interchange fees only for large debit card issuers, while the European Union and Australia regulate interchange fees for all debit and credit card issuers (Hayashi and Maniff 2019). In the United States, the average interchange fee for a credit card transaction is about 2 percent of the transaction value, while the average interchange fee for a debit card transaction is about 0.6 percent of the transaction value for regulated card issuers and 1.15 percent of the transaction value for exempt issuers.3 In contrast, in the European Union, interchange fees are capped at 0.3 percent of the transaction
value for credit cards and 0.2 percent for debit cards. In Australia, interchange fees are capped at 0.8 percent of the transaction value for credit cards and 0.2 percent for debit cards, though card networks also face additional caps.\(^4\) As a result, even a small fraud rate difference affects card issuers’ bottom line in the European Union and Australia, giving issuers a strong incentive to mitigate fraud.

Fourth, even strong incentives—for example, shifting the financial liability for payment fraud from card issuers to merchants—may not be sufficient to overcome some coordination challenges. Although card networks have been using liability shifts to incentivize parties to adopt fraud mitigation tools, such as EMV chip technology and EMV-3D Secure, the liability shift alone may not provide sufficient incentives. For instance, in the United States, liability for fraudulent transactions at fuel pumps not equipped to handle EMV chip cards was supposed to shift from card issuers to convenience stores in October 2017. However, the shift was postponed to October 2020 due to the significant cost of upgrading fuel pumps to support EMV transactions relative to the expected fraud losses convenience stores would avoid by upgrading. It is unclear whether convenience stores, especially smaller ones, will be ready even by the postponed date.\(^5\)

Fifth, card networks themselves may have conflicting interests when it comes to adopting some fraud mitigation tools, such as PINs, in the United States. Although global card networks have mandated PINs for chip card transactions in many other countries, they have not adopted “chip and PIN” as a standard practice in the United States. These networks may want to promote more effective tools than PINs to mitigate fraud in the United States, such as fingerprint or facial recognition on mobile phones. However, global card networks may also want to avoid competing for merchants with domestic debit card networks that require PINs. When cardholders do not use PINs, merchants typically have no choice but to route transactions to global networks; in contrast, when cardholders use PINs, merchants can choose from at least two networks based on the fee charged to merchants. Credit card issuers may also hope to retain or expand their customer base by not adopting PINs; if U.S. consumers consider remembering multiple PINs burdensome, they may limit the number of credit cards they use.
Sixth, consumers in the United States may receive less information about how to mitigate fraud than consumers in other countries. For example, the United Kingdom’s banking and retail industries sponsored a chip-and-PIN advertising campaign that informed consumers about the greater efficacy of PINs in mitigating card-present fraud relative to signatures. In France, the Banque de France has repeatedly communicated with cardholders about their obligations, including keeping PINs safe, protecting card data, and promptly reporting to card issuers any unauthorized transactions or lost or stolen cards. In contrast, U.S. consumers receive little or no information on the efficacy of PINs in mitigating fraud. In fact, U.S. consumers may receive information that encourages them to use more fraud-prone payment methods. For example, some debit card issuers have discouraged their cardholders from using PINs by offering rewards for transactions that use signatures, which carry higher interchange fees.

II. Collecting Data on Fraud Rates

Comparing payment card fraud rates in the United States to those in other countries requires consistent data. However, many countries define fraud in different ways, making direct comparisons of fraud rates challenging. Moreover, some countries do not provide detailed breakdowns of fraud rates by transaction type (for example, in-person versus remote). To account for some of these difficulties, I restrict my comparison to Australia, France, and the United Kingdom. In all three countries, the central bank or a well-established payment organization defines payment fraud and collects detailed fraud statistics.6

Even this restricted sample poses some challenges. For example, the definition of payment fraud is consistent across only three of the four countries. The United States, Australia, and the United Kingdom define payment fraud as a transaction that a third party initiates without the authorization, agreement, or voluntary assistance of the lawful cardholder with the intent to deceive for personal gain. France, however, also includes first-party fraud in their definition. One example of first-party fraud is the authorized cardholder falsely claiming to be defrauded after performing a genuine transaction to purchase goods or services online. Nevertheless, fraud definitions in all four countries share one crucial feature: they do not include attempted fraud that was
prevented before the payment was settled. Thus, only payment fraud that resulted in financial loss, regardless of who incurred such loss, is included in these countries’ fraud statistics.7

All four countries report the overall fraud rate in value—that is, the total value of all fraudulent transactions divided by the total value of all transactions, regardless of transaction channels, card types, and geographic areas. However, the availability of detailed fraud rates differs across countries. The United States and France report fraud rates broken down by transaction type, but Australia and the United Kingdom report only fraud values by transaction type. To calculate fraud rates by transaction type for Australia, I use detailed card transaction data from the RBA, coupled with detailed fraud values reported by the Australian Payments Network (AusPayNet). I cannot calculate detailed fraud rates for the United Kingdom, as detailed card transaction data are not readily available.

Table 1 shows the available fraud rates for different transaction types in all four countries. The availability of different rates varies significantly by country. For example, fraud rates for card-present transactions, which include both ATM and in-person purchase transactions, are available in the United States, Australia, and France, but not in the United Kingdom, which reports only the card-present fraud value. The United States and France divide the card-present fraud rate further into ATM and in-person fraud rates. And the United States subdivides the in-person fraud rate even further based on either authentication technology (chip or no chip) or card verification method (PIN or no PIN). Other countries do not subdivide in-person fraud rates in this way because in-person transactions in these countries typically use both chip and PIN. France and the United Kingdom, however, do report a contactless fraud rate. Card users make contactless transactions by waving or tapping their card at the card reader. Typically, these transactions are limited to small-value transactions and do not require a PIN.8

Although all four countries report remote fraud rates, only France and the United Kingdom report more detailed remote fraud rates. The United Kingdom reports an online fraud rate, and France reports both online and mail-or-telephone order fraud rates. In addition, France reports remote fraud rates for different merchant sectors.
Table 1
Data Availability for Fraud Rates in Value by Country

<table>
<thead>
<tr>
<th>Country</th>
<th>United States</th>
<th>Australia</th>
<th>France</th>
<th>United Kingdom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Transaction types</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Card-present</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>ATM</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>In-person purchase</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Chip versus no chip</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIN versus no PIN</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contactless</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Remote purchase</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Online</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Mail or telephone order</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>By merchant sector</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Card types</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credit versus debit</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transaction or card origin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domestic versus foreign merchants</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Domestic versus foreign cards</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sources: Federal Reserve Board of Governors, AusPayNet, Banque de France, Financial Fraud Action UK, and UK Finance.

The United States distinguishes between debit and credit card fraud, while other countries do not break fraud statistics down by card type. Specifically, the United States reports separate fraud statistics for credit and debit cards and divides debit card fraud further into non-prepaid and prepaid card fraud.

Countries also provide different levels of detail on payment card fraud by card origin. Although all four countries report statistics on fraud conducted with cards issued domestically, Australia, France, and the United Kingdom break domestic card fraud down further based on whether the fraudulent transactions took place at domestic or foreign merchants. In addition, Australia and France report statistics on fraud conducted with foreign-issued cards that are used at domestic merchants.

Finally, on top of the differences in fraud breakdowns shown in Table 1, the cross-country data differ in one other crucial aspect: frequency. The United Kingdom, France, and Australia have collected fraud data every year since 2001, 2002, and 2010, respectively, and all three countries
release the data with a modest delay. For example, 2018 data for all three countries became available within the first seven months of 2019. In contrast, the Federal Reserve System began collecting fraud statistics for the United States in 2012 and only recently released 2015 and 2016 fraud statistics in a report published by the Board of Governors.

III. Comparing Fraud Rates across Countries

To facilitate direct comparisons across all four countries, I restrict my comparison years to 2012, 2015, and 2016—the three years for which detailed U.S. fraud statistics are available. In addition, I focus on fraud conducted with domestic cards, as the United States and the United Kingdom do not report statistics on fraud conducted with foreign cards at domestic merchants. Finally, I focus on fraud rates in value for two reasons. First, detailed fraud rates measured by the number of transactions are unavailable in some countries; and second, the payment industry typically uses fraud rates in value as a benchmark, rather than fraud rates in number.

In-person fraud rates

In 2012, 2015, and 2016, the in-person fraud rate was more than three times higher in the United States than in any other country. Chart 1 compares in-person fraud rates in the United States, Australia, and France with the “contactless fraud rate” in the United Kingdom, the closest measure of in-person fraud available. Although Australia publishes a card-present fraud value, they do not break this value down into in-person and ATM fraud values. Thus, I calculate the highest possible in-person fraud rate for Australia by assuming a zero fraud rate for ATM transactions. Even though Australia’s in-person fraud rates are overstated, the United States’ in-person fraud rates (blue bars) were still higher than the in-person fraud rates in Australia (green bars) by 5 basis points in 2012, 8 basis points in 2015, and 7 basis points in 2016. In addition, the United States’ in-person fraud rates were higher than those in France (orange bars) by 5 basis points in 2012 and by 8 basis points in 2015 and 2016. The United States’ in-person fraud rate was also higher than the contactless fraud rate in the United Kingdom (yellow bars) by 7 basis points in 2016.
Three factors may explain why the in-person fraud rate has been significantly higher in the United States than in other countries. First, the United States has a smaller share of chip transactions in total in-person transactions. EMV migration did not occur in the United States until 2015. In 2016, the first full year after the migration, chip transactions accounted for 23 percent of the value of all in-person transactions. In the western European countries, which include France and the United Kingdom, chip transactions already accounted for 97 percent of the value of all in-person transactions in 2015 (EMVCo 2016). These differences likely contributed to differences in fraud rates, as chip transactions are less likely to be fraudulent overall. The left side of Chart 2 shows that in 2016, the U.S. no-chip fraud rate (blue bar) was 4 basis points higher than the chip fraud rate (green bar) for credit card transactions and 3 basis points higher for debit card transactions.

The second factor that may contribute to the United States’ higher in-person fraud rate is that the United States uses weaker card verification methods with its chip transactions than other countries. In Australia, France, and the United Kingdom, card users provide PINs when making a chip transaction, unless that transaction is contactless. Because only cardholders should know their PINs, these transactions
are less likely to be fraudulent. Indeed, the right side of Chart 2 shows that in 2016, the French fraud rate for contactless transactions (orange bar) was 0.7 basis points higher than for in-person transactions, which include both contactless and chip-and-PIN transactions. Although data on chip-and-PIN transactions alone are not available, the comparison makes clear that contactless transactions are more susceptible to fraud. In contrast, in the United States, the vast majority of credit card chip transactions and some debit card chip transactions are made with no card verification or a weak card verification method, such as a signature. Although some in-person transactions are made with a strong non-PIN card verification method, such as fingerprint verification or facial recognition, those transactions account for a very small proportion of chip transactions. A weak or absent card verification method may partly explain the higher chip fraud rate for U.S. credit cards (the first green bar in Chart 2) than debit cards (the second green bar).

The third factor that may contribute to the United States’ higher in-person fraud rate is that U.S. cardholders are more likely to use credit cards than cardholders in some other countries. Credit card transactions accounted for 44 percent of the value of U.S. in-person transactions in 2016. Although the share was similar in Australia, the share in the United Kingdom was only 28 percent. The equivalent statistic is not available in France, but credit card transactions accounted for only
31 percent of the value of all purchase transactions in 2016. The higher share in the United States may partly explain the higher rate of in-person fraud. The first two sets of bars in Chart 2 show that credit cards carry higher fraud rates than debit cards regardless of whether they use chips. Why credit cards are more prone to fraud than debit cards requires further research; however, the two card types differ notably in both the distribution of transactions between business and consumer cardholders and their shares of card application fraud. In the United States, the share of business transactions was significantly higher in credit card transactions than in debit card transactions (31 versus 9 percent). The share of fraudulent application—perpetrators using stolen identities or false information to obtain a new card and make payments using that card—was also significantly higher for credit cards than debit cards (6.9 versus 0.1 percent).\textsuperscript{15}

**Remote fraud rates**

Unlike in-person fraud rates, the remote fraud rate in the United States has been lower than in Australia and France but higher than in the United Kingdom. Chart 3 shows the remote fraud rates for the United States, Australia, and France in 2012, 2015, and 2016 as well as the e-commerce fraud rate for the United Kingdom in 2015 and 2016.\textsuperscript{16} In 2012, the U.S. remote fraud rate (blue bars) was 27 basis points lower than that of France (orange bars). This gap narrowed to 22 basis points in 2015 and again to 13 basis points in 2016. The U.S. remote fraud rate was also 11 basis points lower than that of Australia (green bars) in 2016. However, the United States has had a higher rate of remote fraud relative to the rate of e-commerce fraud in the United Kingdom (yellow bars). Specifically, the U.S. remote fraud rate was 2 basis points higher in 2015 and 6 basis points higher in 2016.

Two factors may at least partly explain the lower remote fraud rate in the United States relative to Australia and France. First, the vast majority of remote transactions on U.S.-issued cards are made at domestic merchants rather than at foreign merchants. Even if I assume that all U.S. transactions made at foreign merchants were remote transactions, transactions at foreign merchants accounted for less than 6 percent of the value of remote transactions in 2016. In contrast, remote transactions at foreign merchants accounted for 26 percent of the value of all
remote transactions on French-issued cards, more than 13 percent on UK-issued cards, and about 7 percent on Australian-issued cards. These shares likely influence remote fraud rates: although equivalent data for the United States are not available, evidence from the other three countries suggests that remote fraud is significantly more prevalent at foreign merchants than domestic merchants. In 2018, for example, the Australian remote fraud rate was 151 basis points higher at foreign merchants than at domestic merchants.¹⁷

Second, the composition of remote transactions by merchant sector in the United States may differ from other countries. If remote transactions in the United States are more concentrated in merchant sectors with less fraud, such as utilities, the remote fraud rate might be lower than in countries whose remote transactions are more concentrated in higher fraud sectors, such as travel and transportation or online gaming. Although data on merchant composition is not available in the United States, remote fraud varies significantly by merchant sector in France (Banque de France 2019). In addition, Hayashi, Markiewicz,
and Minhas (2018) show that fraud chargeback rates for card-not-present transactions vary significantly by merchant sector in the United States. This rate may be a good proxy for remote fraud rates given that merchants are generally liable for remote fraud. Furthermore, about 20 percent of remote payments in the United States in 2015 were recurring, installment, or other non-purchase payments. These payments may have lower fraud rates than ad hoc purchase transactions because recurring and installment payments require prior contracts between consumers and merchants, such as billers and installment loan providers. These merchants thus know more details about their customers, making them more likely to detect fraudulent transactions.

**Overall fraud rates**

The overall fraud rate, which is the weighted average of in-person, remote, and ATM fraud rates, has been the highest in the United States. Chart 4 shows that the United States had the highest overall fraud rate of all four countries in 2012, 2015, and 2016. The United States’ 11.8 basis points fraud rate in 2016 may in fact be understated: because the U.S. ATM fraud rate is not available for 2016, I assume the ATM fraud rate was zero when constructing the overall fraud rate for that year. Even under this assumption, the gap between the United States and the other three countries appears to have widened over time. For example, the U.S. fraud rate (blue bars) was higher than the rate in the United Kingdom (yellow bars) by 1.1 basis points in 2012, 2.5 basis points in 2015 and at least 3.5 basis points in 2016.

Two main factors may explain the United States’ highest overall fraud rate. First, as discussed previously, the United States has a significantly higher in-person fraud rate than other countries, contributing to its higher overall fraud rate. Second, the United States has a greater share of remote transactions in total card transactions, also likely contributing to its higher overall fraud rate. Although the remote fraud rate is lower in the United States than that in Australia or France, remote transactions are still more prone to fraud than in-person and ATM transactions. Thus, a country with a larger share of remote transactions is more likely to have a higher overall fraud rate.
Conclusion

The United States was one of the last developed countries to migrate to EMV chip technology to mitigate counterfeit card fraud. The United States continues to lag behind some European countries in adopting other fraud-mitigation initiatives, such as chip-and-PIN or 3DS authentication. However, comparing payment card fraud across countries can be challenging: available data vary by country and statistics on U.S. payment card fraud after the EMV migration became available only recently.

I compare in-person, remote, and overall fraud rates in the United States to those in Australia, France, and the United Kingdom, and examine factors explaining the differences. I find that the United States has a significantly higher in-person fraud rate than Australia, France, and the United Kingdom but a lower remote fraud rate than Australia and France. In addition, I find that the United States has the highest overall fraud rate, which is the weighted average of ATM, in-person, and remote fraud rates. A weaker authentication technology (no chip) and a weaker or absent card verification used for many of the in-person transactions—as well as a greater share of credit card transactions for in-person transactions—may explain the United States’ higher in-person
fraud rate. A smaller proportion of remote transactions made at foreign merchants may explain the United States’ lower remote fraud rate. And both the higher in-person fraud rate and greater share of remote transactions in card transactions may explain the United States’ higher overall fraud rate.

Although the overall fraud rate reveals the prevalence of fraud, it may not be a good measure of the effectiveness of fraud mitigation. Fraud rates vary significantly by transaction type, and the composition of transactions across these types varies across countries and may shift from year to year within a country. Detailed fraud rates would help better assess the effectiveness of fraud mitigation. The United States has collected more detailed fraud statistics than some other countries, but it does not break down fraud rates by card verification methods, foreign versus domestic merchants, and business versus consumer card users. Collecting and publicizing these breakdowns may help the U.S. payment industry more effectively monitor and mitigate fraud.
Endnotes

1 3DS is a messaging protocol that strengthens the authorization of online or e-commerce transactions using digital certificates and passwords to authenticate both customer and payment method credentials. The three domains consist of the merchant/acquirer, the issuer, and the payment system. EMV-3D Secure is a new protocol with improved features such as seamless authentication steps, mobile capabilities, and more transaction data. Visa postponed the U.S. activation date for EMV-3D Secure to August 2020, while Mastercard aims to activate the standard in December 2019.

2 Strong customer authentication requires at least two of the following three elements: something the customer knows (such as a password), something the customer has (such as a mobile phone), and something inherent to the customer (such as a fingerprint). Although the effective date of the strong customer authentication requirement was September 14, 2019, the Financial Conduct Authority agreed not to take enforcement action against firms in areas covered by the migration plan until 18 months after the effective date.

3 The interchange fee received by large debit card issuers, defined as issuers with assets of $10 billion or more, is capped at 21 cents per transaction plus 0.05 percent of the transaction value.

4 In addition to the 0.8 percent cap, each credit card network must set interchange fees so that the total value of interchange fees payable on credit card transactions in a year do not exceed 0.5 percent of their total value. An interchange fee for a debit card transaction must not exceed 0.2 percent of the transaction value when the interchange fee is assessed as a percentage of the transaction value and must not exceed 15 cents when the interchange fee is a fixed amount per transaction.

5 The upgrading cost is estimated to be $100,000 to $250,000 per store.

6 The European Central Bank has reported card fraud statistics in the Single European Payments Area (SEPA) (European Central Bank 2018). I exclude the SEPA from the comparison so that I can separately examine France and the United Kingdom, two of the three countries that historically have the highest fraud rates in the SEPA.

7 The United Kingdom reports attempted fraud separately.

8 A supplemental regulation (2018/389) to the revised European Payments Directive limits the value of individual contactless transactions to €50. Cardholders can continue their contactless transactions without using a PIN until their cumulative value of contactless transactions since their last use of a PIN reaches €150 or until they make five consecutive contactless transactions.

9 In France, foreign cards and foreign merchants are further divided into SEPA and non-SEPA cards or merchants.

10 In addition, the Board has reported debit card fraud statistics biennially in its mandatory studies on debit card issuers whose interchange fees are regulated
under Regulation II (Debit Card Interchange Fees and Routing); the most recent study reports the 2017 statistics.

11Fraud statistics involved with foreign cards at domestic merchants have not been collected in the United States and the United Kingdom. However, for domestic merchants and their processors, understanding statistics of fraud involved with foreign cards is important because they could be financially liable for such fraud. In Australia and France, fraud rates of foreign cards are higher than those of domestic cards, especially for remote transactions.

12France reports both in-person and contactless fraud rates, and the latter has been 0.1 to 0.7 basis points higher than the former during the 2015–17 period.

13Although data on chip transactions are unavailable for Australia, the share is likely greater than in the United States because Australia began EMV migration several years earlier.

14In 2017, contactless payments accounted for 3 percent of the value of all in-person transactions in France and 13 percent in the United Kingdom.

15Fraudulent application is the fastest growing fraud type in the United States. This type of fraud may include synthetic identity fraud, in which perpetrators combine fictitious and real information to create new identities to defraud credit card issuers, other financial institutions, government agencies, or individuals. The Federal Reserve Banks (2019) discuss causes and contributing factors of synthetic identity fraud.

16Neither the remote fraud rate nor the e-commerce fraud rate is available for 2012 in the United Kingdom. I use the e-commerce fraud rate for 2015 and 2016 in the United Kingdom because the remote fraud rate is unavailable in those years. In 2018, the remote and e-commerce fraud rates were almost equivalent, suggesting the e-commerce rate may be a good proxy.

17In 2018, the remote fraud rates at domestic and foreign merchants were 14 basis points and 165 basis points in Australia, 17 basis points and 68 basis points in France, and 11 basis points and 25 basis points in the United Kingdom.
References


EMVCo. 2016. “Number of EMV Chip Payment Cards in Global Circulation Increases to 4.8 Billion.” June.


Why Aren’t More People Working in Low- and Moderate-Income Areas?

By Kelly D. Edmiston

Although the U.S. labor market has seen strong growth in recent years, labor market conditions have been weaker in low- and moderate-income (LMI) communities. In particular, residents in LMI communities are much less likely to work than residents in higher-income (non-LMI) communities. As of 2017, 35 percent of residents in LMI communities age 18–64 were not working compared with 24.9 percent in non-LMI communities.

In this article, I use a formal text analysis of a unique set of survey comments to examine prominent obstacles to working, and compare the prevalence of these obstacles, or “employment barriers,” in LMI and non-LMI communities. I find that lower educational attainment and lack of access to transportation and childcare are among the most prominent barriers to employment, and these problems are especially prevalent in LMI communities. Although public assistance, disabilities, and chronic health conditions are considerably more prevalent in LMI communities, they are not especially prominent barriers in the text analysis.

Section I documents the difference in employment rates between LMI and non-LMI communities, showing persistent gaps that are increasing over time. Section II conducts a formal text analysis of survey comments to identify the most prominent barriers to employment. Section III compares statistics on the prevalence of these employment

Kelly D. Edmiston is a senior economist at the Federal Reserve Bank of Kansas City. This article is on the bank’s website at www.KansasCityFed.org
barriers and finds that barriers to work are much more prevalent in LMI communities than non-LMI communities.

I. The Employment Share of the Working-Age Population

To measure differences in employment between LMI and non-LMI communities, I consider only working-age individuals (18–64) and define “communities” as census tracts. Restricting my analysis to the 18–64 population excludes those who are weakly attached to the labor force, such as full-time students and retirees. Defining communities as census tracts (hereafter, “tracts”) allows me to use residence-based employment measures from the U.S. Census Bureau’s American Community Survey, which also contains a wealth of socioeconomic data on demographics, disabilities, and work histories. LMI tracts have median incomes below 80 percent of area median income and make up roughly one-third of all tracts. Tract-level data are available only as five-year averages, the latest of which cover 2013–17 (hereafter, “the 2017 ACS”).

The primary statistic of interest is the employment-to-population ratio (hereafter, “epop ratio”), which is the share of the 18–64 population that is working. In the 2017 ACS, the epop ratio was about 65 percent in LMI tracts, compared with 75.1 percent in non-LMI tracts—a gap of 10.1 percentage points.

Although the epop ratio provides a good aggregate measure of labor market differences in these communities, it does not differentiate between individuals who are not working but actively seeking work (“unemployed”) and individuals who are neither working nor seeking work (“not participating in the labor force”). Quantifying the relative contributions of unemployment and labor force nonparticipation to differences in epop ratios is important because some employment barriers are more likely to affect individuals when looking for a job (such as a criminal conviction), while other barriers may prevent an individual from working altogether (such as a severe disability).

Separating these contributions reveals that differences in labor force nonparticipation explain about three-quarters of the gap in epop ratios between LMI and non-LMI communities, while differences in unemployment explain only one-quarter. Chart 1 shows the epop ratios in LMI and non-LMI communities, the gap between the two, and the
components of the epop ratios from the 2009, 2012, and 2017 ACS. The first set of bars show the epop ratio and associated gap for each ACS period. The second set of bars show that in the 2017 ACS, 6.9 percent of the working-age population in LMI tracts was unemployed compared with 4.4 percent in non-LMI tracts. Consequently, unemployment explains only 2.5 percentage points (6.9 – 4.4) of the total 10.1 percentage point gap in epop ratios between LMI and non-LMI tracts. The third set of bars in Chart 1 shows that in the 2017 ACS, 28.1 percent of the working-age population in LMI tracts did not participate in the labor market compared with 20.5 percent in non-LMI tracts. Thus, nonparticipation explains about 7.6 percentage points (28.1 – 20.5) of the total 10.1 percentage point gap.

Chart 1 also shows that the disparity in labor market outcomes between LMI tracts and non-LMI tracts is persistent. In particular, the gap in the epop ratio is sizeable in all three periods, widening slightly after the Great Recession. Although epop ratios declined for all income groups between the 2009 ACS and 2012 ACS, the decline was somewhat steeper for LMI tracts. In non-LMI tracts, the epop ratio fell by 1.5 percentage points, from 75.0 percent to 73.5 percent. In LMI tracts, the epop ratio declined by 3.3 percentage points, from 66.2 percent to 62.9 percent.
percent. As a result, the gap in the epop ratio between non-LMI tracts and LMI tracts widened from 8.8 percentage points to 10.6 percentage points (after rounding).

The epop ratio recovered for all income groups by the 2017 ACS, but the gap between non-LMI and LMI tracts remained elevated. Although the employment gap narrowed slightly between the 2012 and 2017 ACS, the gap remained 1.3 percentage points higher than in the 2009 ACS. While changes in the epop ratio may be cyclical, the gap between non-LMI and LMI tracts was substantial throughout the business cycle. Specifically, the widening gap in labor force nonparticipation accounted for about 85 percent of the total increase in the employment gap between the 2009 ACS and the 2017 ACS.

II. Identifying Prominent Barriers to Work in Low- and Moderate-Income Areas

The persistent, widening gap in employment between LMI and non-LMI tracts suggests that LMI tracts may face structural barriers to work. To identify potential barriers to work in these communities, I use a unique data set of 258 comments garnered from respondents to the Federal Reserve Bank of Kansas City’s LMI Survey. The LMI Survey is distributed twice yearly to community organizations that work directly and regularly with the LMI population or in LMI communities. The survey uses community organizations as proxies for LMI individuals because surveying LMI individuals on a regular basis can be difficult (Edmiston 2018).

The LMI Survey asks respondents whether economic conditions for the LMI population—including job availability, housing availability, and access to credit—are better, worse, or about the same as the previous quarter. Each of these questions includes a comment box so that respondents can provide further details. In addition, some surveys ask special questions beyond the standard set; the January 2018 survey asked about factors that keep men and women in LMI communities from working.

I use text analysis algorithms on responses to this special survey question to identify common barriers to work. The text analysis is based on natural language processing, which allows computers to understand, interpret, and manipulate human language by applying
a numeric structure to text-based data. I first identify the fundamental argument(s) in each comment and summarize them in a single text (see Appendix A for details on the process, called “latent semantic analysis”).

I then develop a set of terms that match related words in the summarized text. For example, I combine any word related to children, childcare, daycare, parenting, or family responsibilities into the single term “childcare/family.” Similarly, I combine words related to substance abuse and criminal history—both background issues affecting employability—into the term “crime/drugs,” and combine words related to education and training into “ed/training” because of their similar objectives. Aggregating terms in this way ensures that the prominence of a barrier to work is not lost in the many word forms used to describe it.

A word cloud provides a clear way to illustrate the broad themes (terms) as well as the frequency of these themes. Figure 1 shows a word cloud created by feeding the fully prepped text—largely, a long list of terms—to an algorithm. The larger the size of the term in the word cloud, the more frequently the term appears in the text corpus.

Based on the word cloud, the most prominent themes are “jobs,” “qualifications,” and “ed/training.” In the survey, references to “jobs” or related words usually referred to the availability of jobs, though another factor that may have influenced its top billing was the occasional use of “job” as a modifier, as in “job skills.” References to “qualifications” and related words usually addressed inadequate skills for available jobs, while references to education and training usually articulated a need for more access or better quality.

The next most prominent themes, “transportation” and “childcare/family,” could be considered the most direct barriers to employment—individuals cannot work at all without some way of getting to the workplace, and childcare is a necessity for working parents. “Crime/drugs” was the next most prominent theme. Both criminal convictions and substance abuse are “check-the-box” barriers, meaning that simply having a criminal record or failing a drug test often will immediately disqualify an applicant for a job.

“Pay” was the next most prominent theme, though it is difficult to discuss outside of the context of “jobs.” Comments on pay were often about the general need for higher pay, but respondents also mentioned low pay as a disincentive to working. Both “public assistance” and
“housing,” the next most prominent themes, commonly occur along with other barriers. For example, people with disabilities or minor children in the home are more likely to receive public assistance. Likewise, those with criminal convictions may have more difficulty getting approved for housing. Although “disability,” “health,” and “mental,” indicating mental health, were less prominent themes in the word cloud, they are pervasive problems in LMI communities (Barr 2019; Marmot 2002).

Table 1 ranks the most prominent themes drawn from the text analysis and represented in the word cloud. My analysis excludes some words in the table. I exclude “affordable,” for example, because it was used exclusively as a modifier for other terms, such as “affordable childcare” or “affordable housing,” and has little meaning out of context. I also exclude “motivation,” which sometimes referred to individuals being motivated to work but being unsuccessful, but more commonly referred to an individual lacking the motivation to seek a job. Little can be done with this concept in terms of a quantitative analysis, as distinguishing the context in which the word was used is unfeasible and I am not able to measure motivation. Finally, I exclude “government” or similar terms that were explicitly political or that referred to funding available to the

Figure 1
Word Cloud of Common Terms Used in LMI Survey Responses

Note: The size of the term is proportional to its frequency in the analyzed text.
Source: Author’s calculations.
Table 1
Major Themes from the Text Analysis and Associated Words

<table>
<thead>
<tr>
<th>Rank</th>
<th>Term</th>
<th>Examples of associated words (not comprehensive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jobs</td>
<td>Jobs, work, employment, unemployment, layoffs, positions</td>
</tr>
<tr>
<td>2</td>
<td>Qualifications</td>
<td>Qualifications, qualify, qualified, skills, skill sets, mismatch, employable, requirements, credentials, unskilled, marketable, standards</td>
</tr>
<tr>
<td>3</td>
<td>Ed/training</td>
<td>Education, training, workforce development, mentoring, literacy, GED, educate, high school, graduate, educational, train, degree(s)</td>
</tr>
<tr>
<td>4</td>
<td>Transportation</td>
<td>Transportation, transit, car(s), proximity, close</td>
</tr>
<tr>
<td>5</td>
<td>Childcare/family</td>
<td>Childcare, daycare, children, family, families, kids, parent(s), pre-school</td>
</tr>
<tr>
<td>6</td>
<td>Crime/drugs</td>
<td>Criminal record(s), criminal history, felony, conviction(s), ex-offender, drug(s), substance, alcohol, addiction, background issues</td>
</tr>
<tr>
<td>7</td>
<td>Pay</td>
<td>Pay, paying, wage(s), salaries</td>
</tr>
<tr>
<td>8</td>
<td>Public assistance</td>
<td>Government benefits, benefits, assistance, SSI, welfare, dependence</td>
</tr>
<tr>
<td>9</td>
<td>Housing</td>
<td>Housing, homeless(ness), home</td>
</tr>
<tr>
<td>10</td>
<td>Mental</td>
<td>Mental health, mental illness, mental, mentally, low functioning</td>
</tr>
<tr>
<td>11</td>
<td>Motivation</td>
<td>Want to work, unwilling to work, initiative, willingness, work ethic</td>
</tr>
<tr>
<td>12</td>
<td>Government</td>
<td>Government, federal, state, politics, political, city, funds, resources</td>
</tr>
<tr>
<td>13</td>
<td>Health</td>
<td>Health, medical, physical illness, illness, sick, healthy</td>
</tr>
<tr>
<td>14</td>
<td>Disability</td>
<td>Disability, disabilities, disabled, impairment</td>
</tr>
<tr>
<td>15</td>
<td>Affordable</td>
<td>Affordable, afford, cost</td>
</tr>
</tbody>
</table>

Notes: The “motivation,” “government,” and “affordable” themes are not specifically analyzed in the text. Associated words are identified through lemmatization, a linguistic process that groups together the inflected forms of a word (for example, “run” and “ran”) for analysis as a single item.

To better understand the importance of the employment barriers identified in the text analysis, I compare the prevalence of these barriers in LMI tracts with their prevalence in non-LMI tracts. These comparisons require that I transform qualitative responses from the sentiment analysis into quantitative measures. Thus, for each barrier, I locate or construct a quantitative indicator, or proxy. As an example, I use the share of households in the tract without access to a vehicle as a quantitative indicator for “transportation.” For robustness, I consider multiple indicators for most terms based on how well the quantitative indicator represents the qualitative sentiment and on the availability of data.
Table 2 provides statistics for the indicators used to measure each barrier. Column 1 shows the mean value of these indicators in LMI tracts, column 2 shows the mean value of these indicators in non-LMI tracts, and column 3 shows the difference in the means of the indicators between LMI and non-LMI tracts. The difference in means for every indicator is statistically significant, meaning I can conclude with meaningful certainty that the true difference in the barrier’s prevalence between LMI tracts and non-LMI tracts is not zero. In the vast majority of cases, barriers are more prevalent and severe in LMI tracts. A statistically significant difference is not necessarily economically significant, however. To gauge economic significance, I also report the ratio of the difference in means to the mean in non-LMI tracts in column 4.

Although Table 2 includes multiple indicators for each barrier, I examine only a few indicators in detail in the subsequent analysis for tractability. The indicators in Table 2 that are not discussed serve as “robustness checks,” providing additional support for the conclusions drawn. See Appendix B for the data sources and Appendix C for details on the construction of each indicator.

Jobs and pay

Although “jobs” was the most common barrier cited in the LMI survey, the context of job-related comments varied widely. Some comments implied plenty of jobs were available, while others implied an insufficient number of jobs were available. To draw conclusions from these conflicting assessments, I use the LMI Job Availability Index, which tracks the diffusion of survey responses to a question about the availability of jobs in LMI communities over time. Any index value above 100 (neutral) means that more survey respondents stated jobs were more available than stated jobs were less available. Chart 2 shows the index relative to the previous year (blue line) and quarter (green line) alongside expectations for the following quarter (orange line). All three indexes were above neutral in every quarter after 2012, which means the balance of survey opinion has been that jobs are plentiful in LMI-relevant sectors—or at least increasingly so.

Measuring job availability in LMI tracts relative to non-LMI tracts is challenging because residents in these tracts essentially face the same geographic labor market. Most people do not live and work in the same
### Table 2

**Barriers to Employment in LMI and Non-LMI Census Tracts**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (LMI)</th>
<th>Mean (non-LMI)</th>
<th>Difference in means</th>
<th>Percent difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Jobs and pay</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workers in tract / residents (W/P)</td>
<td>0.73</td>
<td>1.00</td>
<td>-0.268***</td>
<td>-26.7</td>
</tr>
<tr>
<td>Health workers / residents</td>
<td>0.13</td>
<td>0.16</td>
<td>-0.031*</td>
<td>-19.5</td>
</tr>
<tr>
<td>Retail workers / residents</td>
<td>0.09</td>
<td>0.12</td>
<td>-0.032***</td>
<td>-25.5</td>
</tr>
<tr>
<td>Accommodations and food service workers / residents</td>
<td>0.08</td>
<td>0.10</td>
<td>-0.017***</td>
<td>-17.7</td>
</tr>
<tr>
<td>Median earnings (age 16+)</td>
<td>$41,977</td>
<td>$58,375</td>
<td>-$16,398***</td>
<td>-28.1</td>
</tr>
<tr>
<td><strong>Qualifications</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent age 25+ with less than 9th grade</td>
<td>8.2</td>
<td>3.7</td>
<td>4.5***</td>
<td>120.3</td>
</tr>
<tr>
<td>Percent age 25+ with no high school diploma</td>
<td>19.4</td>
<td>9.6</td>
<td>9.8***</td>
<td>102.2</td>
</tr>
<tr>
<td>Percent age 25+ with associate degree or some college</td>
<td>28.4</td>
<td>28.7</td>
<td>-0.2***</td>
<td>-0.8</td>
</tr>
<tr>
<td>Percent age 25+ with bachelor’s degree or more</td>
<td>20.1</td>
<td>33.9</td>
<td>-13.9***</td>
<td>-40.9</td>
</tr>
<tr>
<td>Percent age 18–64 with no work in past 12 months</td>
<td>30.2</td>
<td>21.7</td>
<td>8.6***</td>
<td>39.5</td>
</tr>
<tr>
<td>Percent residents age 15–29</td>
<td>23.0</td>
<td>18.7</td>
<td>4.3***</td>
<td>23.1</td>
</tr>
<tr>
<td><strong>Ed/training</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miles to closest training center</td>
<td>6.8</td>
<td>7.9</td>
<td>-1.0***</td>
<td>-13.2</td>
</tr>
<tr>
<td>Density of training centers (per square mile)</td>
<td>0.07</td>
<td>0.04</td>
<td>0.023***</td>
<td>53.3</td>
</tr>
<tr>
<td><strong>Transportation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent households with no vehicle</td>
<td>16.5</td>
<td>6.6</td>
<td>9.9***</td>
<td>149.7</td>
</tr>
<tr>
<td>Percent age 18–64 self-employed</td>
<td>6.7</td>
<td>6.9</td>
<td>-0.1***</td>
<td>-1.7</td>
</tr>
<tr>
<td><strong>Childcare/family</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Childcare facilities per square mile</td>
<td>1.8</td>
<td>1.1</td>
<td>0.7***</td>
<td>59.4</td>
</tr>
<tr>
<td>Average family size</td>
<td>3.3</td>
<td>3.1</td>
<td>0.2***</td>
<td>6.6</td>
</tr>
<tr>
<td>Percent households with children age &lt;18</td>
<td>26.6</td>
<td>26.9</td>
<td>-0.3***</td>
<td>-1.1</td>
</tr>
<tr>
<td>Percent female-headed households with children age &lt;18</td>
<td>10.0</td>
<td>5.0</td>
<td>5.0***</td>
<td>99.7</td>
</tr>
<tr>
<td>Percent male-headed households with children age &lt;18</td>
<td>2.8</td>
<td>2.0</td>
<td>0.7***</td>
<td>34.3</td>
</tr>
<tr>
<td>Childcare costs (U.S. average)/earnings (percent)</td>
<td>24.1</td>
<td>16.3</td>
<td>7.8***</td>
<td>47.9</td>
</tr>
<tr>
<td><strong>Crime/drugs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crime rate (annual per 10,000 people)</td>
<td>93.7</td>
<td>57.2</td>
<td>36.5***</td>
<td>63.7</td>
</tr>
<tr>
<td>Drug deaths (annual per 10,000 people)</td>
<td>2.38</td>
<td>2.33</td>
<td>0.05***</td>
<td>2.0</td>
</tr>
<tr>
<td>Alcohol deaths (annual per 10,000 people)</td>
<td>1.18</td>
<td>1.14</td>
<td>0.04***</td>
<td>3.7</td>
</tr>
<tr>
<td>Annual opioid prescription / 100 residents</td>
<td>61.1</td>
<td>60.6</td>
<td>0.5*</td>
<td>0.8</td>
</tr>
<tr>
<td><strong>Housing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent households renters</td>
<td>49.5</td>
<td>28.0</td>
<td>21.6***</td>
<td>77.2</td>
</tr>
<tr>
<td>Percent households paying &gt;35 percent of income in rent</td>
<td>42.6</td>
<td>31.8</td>
<td>10.8***</td>
<td>34.1</td>
</tr>
<tr>
<td>Percent households in different house in same county</td>
<td>10.3</td>
<td>7.0</td>
<td>3.3***</td>
<td>46.2</td>
</tr>
<tr>
<td>Percent households with more people than rooms</td>
<td>4.9</td>
<td>2.2</td>
<td>2.7***</td>
<td>124.3</td>
</tr>
<tr>
<td>Variable</td>
<td>Mean (LMI)</td>
<td>Mean (non-LMI)</td>
<td>Difference in means</td>
<td>Percent difference</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>------------</td>
<td>----------------</td>
<td>---------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Homeless per 100 residents</td>
<td>0.19</td>
<td>0.16</td>
<td>0.03***</td>
<td>16.3</td>
</tr>
<tr>
<td>Homeless per square mile</td>
<td>10.54</td>
<td>6.52</td>
<td>4.02***</td>
<td>61.7</td>
</tr>
<tr>
<td>Chronically homeless per 100 residents</td>
<td>0.03</td>
<td>0.02</td>
<td>0.005***</td>
<td>19.7</td>
</tr>
<tr>
<td>Chronically homeless per square mile</td>
<td>0.9</td>
<td>0.7</td>
<td>0.2***</td>
<td>28.2</td>
</tr>
<tr>
<td>Disability/mental/health</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent age 18–64 with disability</td>
<td>14.6</td>
<td>9.9</td>
<td>4.7***</td>
<td>47.1</td>
</tr>
<tr>
<td>Percent age 18–64 with ambulatory disability</td>
<td>7.7</td>
<td>4.8</td>
<td>3.0***</td>
<td>62.1</td>
</tr>
<tr>
<td>Percent age 18–64 with cognitive disability</td>
<td>6.5</td>
<td>4.1</td>
<td>2.5***</td>
<td>47.1</td>
</tr>
<tr>
<td>Age-adjusted mortality rate (annual, per 100,000)</td>
<td>77.8</td>
<td>74.9</td>
<td>2.9***</td>
<td>3.9</td>
</tr>
<tr>
<td>Public assistance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent households receiving public assistance</td>
<td>24.1</td>
<td>9.4</td>
<td>14.6***</td>
<td>155.1</td>
</tr>
<tr>
<td>Percent households receiving TANF</td>
<td>4.3</td>
<td>1.9</td>
<td>2.4***</td>
<td>125.9</td>
</tr>
<tr>
<td>Percent households receiving SNAP</td>
<td>23.3</td>
<td>8.8</td>
<td>14.4***</td>
<td>163.7</td>
</tr>
<tr>
<td>Percent households receiving SSI</td>
<td>8.9</td>
<td>4.3</td>
<td>4.5***</td>
<td>104.5</td>
</tr>
</tbody>
</table>

* Significant at the 10 percent level
** Significant at the 5 percent level
*** Significant at the 1 percent level

Notes: The difference in means may not align with the reported means due to rounding. F-fold statistics reject variance equality for virtually all variables, where $F_{ij} = \max(\sigma_{ij}^2/\sigma_{jj}^2) / \min(\sigma_{ij}^2/\sigma_{jj}^2)$ and $\sigma_{ij}$ is the row $i$, column $j$ element of the covariance matrix. Therefore, t-statistics (not reported) use Satterthwaite’s approximation for degrees of freedom. Statistical significance is determined using Cochran p-values.

Chart 2

**LMI Job Availability Index**

Index: 100 = neutral

Note: The survey asks respondents to assess conditions relative to the same period in the previous year, conditions relative to the previous quarter, and for their expectations for the following quarter relative to the current quarter.

Source: Federal Reserve Bank of Kansas City.
tract. The average commute distances for the 69 largest U.S. metropolitan areas range from 5.0 to 12.8 miles (Kneebone and Holmes 2015). The geographic labor markets, which I define as circles with radii equal to average commuting distances, would therefore include many tracts. For example, using a representative LMI neighborhood in Kansas City, Missouri, where the average commute is 8.9 miles, I measure a labor market area consisting of 213 tracts.

An analysis of more localized labor markets may offer some insight, at least to the extent that there are benefits to having jobs nearby. To measure job opportunities in a more localized labor market, I compare the number of people who work in a tract with the number of people who live in the tract. The premise underlying this measure is that the number of workers in a tract is a reasonable (albeit imperfect) indicator of the number of jobs available in the tract. If there are more workers in the tract, I presume there are more job opportunities in the tract. Importantly, workers/residents is different from residents with jobs/residents. Most of those who work in a tract live in a different tract. Likewise, most of those who live in a tract work in a different tract. My calculation of workers/residents shows that on average, people who live in non-LMI tracts have more nearby job opportunities. Specifically, Table 2 shows that LMI tracts have 0.73 workers per resident compared with 1 worker per resident in non-LMI tracts.6

Much like job availability, pay is a challenging indicator to evaluate because, again, residents in LMI and non-LMI tracts face essentially the same geographic labor market. The importance of pay as a barrier to employment depends on how responsive potential workers are to different rates of pay in deciding whether to work. Research suggests, for example, that marginally higher pay has little effect on this decision; most people would need to achieve a certain pay threshold to be induced to work (McClelland and Mok 2012). However, a nontrivial share of survey comments asserted that prevailing wages are disincentives to work, suggesting their constituents would need substantially higher pay. Indeed, specific comments mentioned that for many, working does not seem worthwhile when the pay is insufficient to sustain them or their families.

If self-sufficiency is required to make work worthwhile, then pay is likely a more significant barrier to work in LMI tracts than in non-LMI
tracts. However, the difference in pay results from the types of jobs that are attainable based on qualifications, rather than wage differentials between distinct labor markets faced by those in LMI and non-LMI tracts.\(^7\)

Although the text analysis suggests that job availability and pay may be critical factors in decisions about working for the LMI population, disparities between LMI and non-LMI tracts likely arise from differences in the types of jobs for which residents qualify and the compensation those jobs offer, not geographic differentials.

**Qualifications, education, and training**

Education and training, along with work experience, are unquestionably advantages in the labor market. Labor market statistics clearly document returns to educational attainment in the form of lower unemployment rates and higher average earnings. However, the type of education received is also important. Hanushek and others (2016) suggest that while specific skills gained in vocational training may ease the transition to a first job, the specificity of the training may make workers less adaptable for future work compared with those with a more general education.

To help quantify the importance of education differentials as barriers to employment, I compare educational attainment among individuals age 25 and older between LMI and non-LMI tracts. Table 2 shows that 19.4 percent of individuals in LMI tracts have not earned a high school diploma or equivalent, compared with only 9.6 percent of individuals in non-LMI tracts. The rates for those with an associate degree or “some college” are similar in LMI and non-LMI tracts, potentially reflecting a greater share of LMI individuals with vocational training. However, substantially fewer individuals in LMI communities have a bachelor’s degree. Specifically, 20.1 percent of individuals in LMI tracts have a bachelor’s degree or higher compared with 33.9 percent in non-LMI tracts.

Skills come not only from formal education and training but also from experience. Residents in LMI tracts typically have less experience than residents in non-LMI tracts, as indicated by the percentage of the population age 15–29 (presuming young people have less job experience) and the share who have not worked at all in the preceding 12 months (presuming skill atrophy or obsolescence).
Together, these statistics suggest that lacking qualifications is a substantial impediment to work. As the second most frequently mentioned employment barrier in LMI Survey comments, a lack of qualifications appears to be a widespread problem, compounded by large gaps in education and experience between LMI and non-LMI tracts.

The need for additional education and training opportunities was unsurprisingly a common refrain in LMI Survey comments. Because the majority of tracts contain no facility, I calculate the distance to the closest facility. The average distance to workforce training is only about a mile shorter, on average, in LMI tracts (6.8 miles) than in non-LMI tracts (7.9 miles). By these measures, the proximity of training opportunities does not appear to differ substantially in LMI and non-LMI tracts.

**Transportation**

Most people work in a different tract than the one in which they live, and commuting distance may be a significant barrier to work for many people. The average commute range of 5.0 to 12.8 miles reported by Kneebone and Holmes (2015) suggest significant hurdles for those with few transportation options. The greatest transportation barrier is likely lack of access to a vehicle (Baum 2009; Blumberg and Pierce 2014; King, Smart, and Manville 2019). Vehicle access is also arguably the most straightforward transportation barrier to measure, as the ACS reports the share of households without a vehicle.

The difference in vehicle access between LMI tracts and non-LMI tracts is quite stark. Table 2 shows that 16.5 percent of households in LMI tracts do not have access to a vehicle. By contrast, only 6.6 percent of households in non-LMI tracts lack access to a vehicle.

Households without a vehicle may prefer to work close to home, but data suggest this option is not as viable in LMI tracts as in non-LMI tracts. As noted in the discussion of job availability, the number of workers per resident is lower in LMI tracts. In addition, self-employment rates are also lower in LMI tracts. As with qualifications, a lack of transportation is both a pervasive employment barrier—as indicated by its ranking in the text analysis—and considerably more prevalent in LMI communities than non-LMI communities.
Childcare and family issues

Childcare availability and cost are frequent concerns for working parents and for parents who would like to work. In a recent poll, 70 percent of respondents reported location to be one of the most important factors they consider when choosing a provider (Dodge-Ostendorf and others 2019). On the cost side, Powell (2002) provides causal statistical evidence that the high cost of childcare reduces the probability of working. The cost of childcare may lead a parent to reasonably question whether working is financially worthwhile.

Although there is not an ideal measure of childcare costs on a geographic basis, I measure physical access to childcare by calculating the density of childcare establishments—that is, the number of childcare facilities per square mile in a tract. Childcare facilities are generally more accessible in LMI tracts, which have 1.8 childcare facilities per square mile, than in non-LMI tracts, which have 1.1 facilities per square mile (Table 2). Thus, physical access to childcare alone does not appear to be a greater barrier in LMI areas, though the cost of nearby facilities could alter the calculus.

Even if greater competition (as measured by density) effectively reduced costs, childcare likely would still be much less affordable in LMI tracts, where income is much lower. On average, childcare costs $8,606 annually in the United States, though there is substantial geographic variation (Child Care Aware 2017). Using this national average, childcare costs are 24.1 percent of median earnings in LMI tracts, compared with 16.3 percent of median earnings in non-LMI tracts.

Furthermore, working families in LMI tracts may have a greater need for childcare. The share of households with minor children is roughly the same across tracts, but households with minor children are twice as likely to be headed by a single mother in LMI tracts and therefore lack a spousal childcare option or spousal income support for childcare (single father households also are more common in LMI tracts, but rarer in general).8

Crime and substance abuse

People with criminal convictions have a significant disadvantage in finding employment compared with those without criminal convictions. Pager (2003) provides causal evidence that simply having a
criminal record reduces employment opportunities irrespective of potential delays in education and experience due to incarceration or personality traits that may be common among those with criminal records but separate from their criminal behavior. References to this problem were prominent in LMI Survey comments and also pervasive in focus-group discussions with non-working LMI individuals recently hosted by the Federal Reserve Banks of Chicago and Kansas City. The increasing use of criminal background checks in employment pre-screening makes the problem even more formidable (Blumstein and Nakamura 2009).

There is no practical way to determine how many people in a geographic area have criminal records, making this barrier especially difficult to evaluate. However, Wiles and Costello (2000) find that offenders are most likely to commit crimes near their homes, which suggests crime rates are correlated with the presence of offenders. I estimate tract-level crime rates for 2008–12 and use them as a proxy for the prevalence of individuals with criminal convictions in 2013–17.

My estimates of crime rates are dramatically higher for LMI tracts than non-LMI tracts. From 2008 to 2012, LMI tracts had 93.7 crimes per 10,000 residents, compared with 57.2 crimes per 10,000 residents in non-LMI tracts (Table 2). If Wiles and Costello (2000) are correct that criminals tend to commit their crimes close to home—and if a higher crime rate in 2008–12 is associated with a higher percentage of the population having criminal convictions in 2013–17—then criminal convictions may be more prevalent barriers to employment in LMI tracts than non-LMI tracts.

In contrast, substance abuse may not be a substantially more prevalent barrier in LMI communities. In 2017, nonintentional drug-related deaths averaged 2.4 per 10,000 residents in LMI tracts and 2.3 per 10,000 residents in non-LMI tracts. Alcohol-related deaths were only modestly different, averaging 1.2 and 1.1 deaths per 10,000 residents in LMI tracts and non-LMI tracts, respectively.

Opioid use also appears to be similar in LMI and non-LMI communities. Opioid prescriptions are known to reduce labor force participation, making them an especially useful indicator of the connection between substance abuse and employment rates (Aliprantis and others 2019; Krueger 2017). The data, which are for 2015, show very little difference between LMI tracts, where 61.1 opioid prescriptions were
written annually per 100 residents, and non-LMI tracts, where 60.6 prescriptions were written annually. While substance use may be an impediment to employment, the evidence does not point to substantially higher rates of substance abuse in LMI tracts. The differences are statistically significant but negligible in economic significance.

### Housing instability

A lack of secure and stable housing may be a barrier to employment for multiple reasons. Housing instability can upset social ties or prevent the formation of social ties, which provide an important source of information about job opportunities and may help workers address sudden needs for transportation or childcare (Briggs 1998; Calvo-Armengol and Jackson 2004). In addition, housing instability consumes time and focus and can induce significant stress, making it more difficult to find or retain a job (Manzo and others 2008).

I evaluate housing instability using a variety of data from the ACS and consistently find greater instability in LMI tracts than non-LMI tracts. Perhaps most tellingly, households in LMI tracts are 34 percent more likely than households in non-LMI tracts to devote over 35 percent of their gross income to rent. Allocating such a large share of income to rent increases the likelihood that a household will be unable to make rent payments (Desmond and Shollenberger 2015). In addition, households in LMI tracts are more likely to live in renter-occupied units (49.5 percent versus 28.0 percent in non-LMI tracts), live in a different house in the same county than the year before (10.3 percent versus 7.0 percent), or live in housing units with more residents than rooms (4.9 percent versus 2.2 percent). By every measure, housing appears to be more unstable in LMI tracts than in non-LMI tracts.

The extreme side of housing instability is, of course, homelessness. Homelessness can present unique barriers to employment, such as shelter policies that limit the ability to work odd hours (Poremski and others 2016). However, data on homelessness are unsurprisingly difficult to obtain, given that homeless people do not have a stable physical address. I use counts from the U.S. Department of Housing and Urban Development (HUD)’s Continuum of Care (CoC) Program to estimate the number of homeless people per square mile (homeless density) and the number of homeless people per 100 residents (homeless rate).
Homeless density is 10.54 per square mile in LMI tracts and 6.52 per square mile in non-LMI tracts. The homeless rate is 0.19 per hundred residents in LMI tracts and 0.16 per hundred residents in non-LMI tracts. Thus, homelessness appears to be moderately more prevalent in LMI tracts.

**Disabilities and mental and physical health**

Disabilities and poor health—both physical and mental—are direct barriers to work in that they put limits on what a worker is effectively able to accomplish. Some disabilities or illnesses may prevent workers from doing certain jobs at all.

Data from the ACS show that people in LMI tracts are much more likely to have disabilities than people in non-LMI tracts. The ACS measures the presence of any disability as a “yes” response to at least one of its six disability questions. In LMI tracts, 14.6 percent of the working-age population report having some disability, compared with 9.9 percent of the working-age population in non-LMI tracts.

The ACS also differentiates between cognitive and ambulatory disabilities, which helps capture the distinction made in LMI Survey comments between mental and physical health. Cognitive disabilities are more common among residents in LMI tracts (6.5 percent) than non-LMI tracts (4.1 percent). These disabilities may make finding and retaining a job more difficult. Among the most frequent work problems for those with cognitive disabilities are lack of motivation, side effects from medication, substance abuse, low self-confidence, stigma, treatment issues, and difficulties in identifying and achieving goals (Secker and others 2001; Bassett, Lloyd, and Bassett 2001). Ambulatory disabilities are also much more common among residents in LMI tracts (7.7 percent) than non-LMI tracts (4.8 percent). One frequent work-related problem for those with ambulatory or other physical disabilities is a limitation on the tasks they are physically able to complete. In addition, a lack of social acceptance by coworkers can keep employees with disabilities from staying in jobs, making their employment less stable (Shier, Graham, and Jones 2009).

Health issues can also lead to less stable employment, and research documents a causal effect of health on employment rates. Wilson (2001) estimates that chronic adult-onset disease explains 10 percent
of nonemployment among those age 35–74 in New Jersey. Zhang and others (2009) find significant causal effects of several chronic diseases on employment. For example, diabetes lowers the probability of employment by about 4 percentage points for men age 18–49 and by 11.5 percentage points for older men.

One general measure of health available for a large number of tracts is the age-adjusted mortality rate, which in 2017 was moderately higher in LMI tracts (77.8) than in non-LMI tracts (74.9). Although data on specific health conditions are also available for multiple years, they are only available for tracts in the nation’s 500 largest cities, which account for about one-third of the U.S. population. Table 3 shows that most specific health conditions are considerably more prevalent in LMI tracts than non-LMI tracts, the exception being cancer. Rates of self-reported physical and mental health are 51 percent and 41 percent higher in LMI tracts. Many unhealthy behaviors correlated with chronic illness are also much more common in LMI tracts. Obesity and smoking are 36 percent and 49 percent more prevalent in LMI tracts than in non-LMI tracts, respectively. An exception is binge-drinking, which is more common in non-LMI tracts.

Although disabilities and poor health were not among the most commonly cited barriers to employment in the LMI Survey, their greater prevalence in LMI communities, along with research showing significantly lower employment among the disabled and chronically ill, suggests they may be important nonetheless.

**Public assistance**

People who do not work—particularly those with disabilities, health problems, and minor children in the home—often receive public assistance, which may discourage working in the future. For example, Maestas, Mullen, and Strand (2013) find that the employment rate of beneficiaries on the margin of entry into the Social Security Disability Insurance (SSDI) program in 2005–06 would have been 28 percentage points higher two years later if they had never received SSDI benefits. But perhaps more importantly, most public assistance programs are structured in a way that discourages recipients from working even in the absence of any income effects.
A highly significant work disincentive built into public assistance programs is the benefit reduction scheme associated with earned income. Benefit reductions are similar to a tax on earned income. For example, at certain levels of income, Supplemental Nutrition Assistance Program (SNAP) benefits are reduced by 30 cents per dollar earned. Housing, childcare, or cash assistance through Temporary Assistance for Needy Families (TANF) also are reduced at some level of earned income. Benefit reduction rates vary widely by state. Some states have exemptions, usually time-limited and capped, but eventually benefit reductions come into play. The Earned Income Tax Credit is a substantial offset to these benefit reductions. Still, Maag and others (2012) document cases of marginal effective tax rates (tax rate on the next dollar inclusive of benefit reduction rates) greater than 100 percent. With such high marginal effective tax rates, beneficiaries may reasonably decide that work or additional work is not worthwhile, especially when they consider costs for childcare and transportation.

A considerably larger share of residents in LMI tracts receive public assistance than in non-LMI tracts. Overall, households in LMI

Table 3
Prevalence of Health Problems in LMI and Non-LMI Census Tracts

<table>
<thead>
<tr>
<th>Health indicator</th>
<th>Mean (LMI) (percent)</th>
<th>Mean (non-LMI) (percent)</th>
<th>Difference in means</th>
<th>Percent difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence coronary heart disease</td>
<td>6.7</td>
<td>5.2</td>
<td>1.5</td>
<td>29.0</td>
</tr>
<tr>
<td>Prevalence poor mental health</td>
<td>15.4</td>
<td>10.9</td>
<td>4.5</td>
<td>41.4</td>
</tr>
<tr>
<td>Prevalence poor physical health</td>
<td>15.5</td>
<td>10.3</td>
<td>5.2</td>
<td>50.6</td>
</tr>
<tr>
<td>Prevalence arthritis</td>
<td>24.3</td>
<td>21.4</td>
<td>2.9</td>
<td>13.5</td>
</tr>
<tr>
<td>Prevalence asthma</td>
<td>10.9</td>
<td>9.0</td>
<td>1.9</td>
<td>21.7</td>
</tr>
<tr>
<td>Prevalence binge drinking</td>
<td>16.1</td>
<td>19.7</td>
<td>-3.6</td>
<td>-18.3</td>
</tr>
<tr>
<td>Prevalence cancer</td>
<td>5.1</td>
<td>6.0</td>
<td>-0.9</td>
<td>-15.7</td>
</tr>
<tr>
<td>Prevalence diabetes</td>
<td>13.3</td>
<td>8.7</td>
<td>4.7</td>
<td>53.9</td>
</tr>
<tr>
<td>Prevalence obesity</td>
<td>35.0</td>
<td>25.8</td>
<td>9.3</td>
<td>36.0</td>
</tr>
<tr>
<td>Prevalence smoking</td>
<td>22.1</td>
<td>14.8</td>
<td>7.2</td>
<td>48.6</td>
</tr>
</tbody>
</table>

Notes: The difference in means is significant at the 1 percent level and may not align with the reported means due to rounding. F-fold statistics reject variance equality for virtually all variables, where $F' = \max(s_{11},s_{22})/\min(s_{11},s_{22})$ and $s_{ij}$ is the row $i$, column $j$ element of the covariance matrix. Therefore, $t$-statistics (not reported) use Satterthwaite’s approximation for degrees of freedom. Statistical significance is determined using Cochran p-values.

Sources: Centers for Disease Control and Prevention and author’s calculations.
tracts receive public assistance at 2.5 times the rate of households in non-LMI tracts. Differences in the rates at which households receive public assistance are similar across programs. The higher rate of Supplemental Security Income (SSI) receipt in LMI tracts results largely from a greater share of residents in LMI tracts providing care for disabled children.

**Summary and Conclusions**

Working-age residents in LMI tracts are less likely to work than working-age residents in non-LMI tracts. The gap in epop rates is quite large—10.1 percentage points in the latest available data—but also persistent. Moreover, in recent years, the gap has been growing, due mostly to differences in the share of the working-age population neither working nor looking for work.

Based on a text analysis of a unique set of survey responses to a question on relatively low employment rates in LMI communities, I identify several potential employment barriers, rank their prominence in the survey comments, and then compare their prevalence in LMI and non-LMI tracts. The analysis suggests that barriers are more prevalent in LMI tracts across the board, though educational attainment, transportation, and childcare are especially prominent and prevalent in LMI tracts. Although barriers such as mental and physical disabilities and poor health did not rank especially high in the survey comments, they are considerably more prevalent in LMI communities, suggesting they may nevertheless warrant close attention.

These results may be useful to agents in the social services sector seeking to allocate resources toward improving LMI employment outcomes. In particular, my analysis suggests that overcoming barriers to education and training, transportation, and childcare may help improve employment in LMI tracts. Restructuring public assistance programs to reduce disincentives for work and improving public health efforts in LMI communities may also help more individuals in LMI tracts enter the workforce.
Appendix A

Latent Semantic Analysis (LSA)

The first step in LSA is to “tokenize” the comments by chopping comments into pieces (often individual words) called tokens. Punctuation is removed, as are “stop words”—extremely common words such as “the” and “a” that would be of little value in understanding the text.

A process called “lemmatization” reduces inflectional forms and derivationally related forms of a word by grouping tokens to a common base word. For example, lemmatization would reduce “run,” “running,” and “ran” to the word “run,” and “am,” “are,” and “is” to the word “be.” The resulting tokens are then encoded as numbers but maintain their association with the sentence or passage from which they originated in the form of a numeric matrix. The matrix is manipulated for use in LSA. Specifically, a term-weighted matrix is created based on the frequencies of words co-occurring, and a singular value decomposition is performed on the resulting matrix. A topic is identified purely on the likelihood of words co-occurring and has no basis in connotation.
## Appendix B

### Data Sources

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment rate (by residence)</td>
<td>2017 ACS 5-Year Estimates, Table 2301</td>
</tr>
<tr>
<td>Unemployment rate (by residence)</td>
<td>2017 ACS 5-Year Estimates, Table 2301</td>
</tr>
<tr>
<td>Labor force participation rate (by residence)</td>
<td>2017 ACS 5-Year Estimates, Table 2301</td>
</tr>
<tr>
<td>Population 18–64</td>
<td>2017 ACS 5-Year Estimates, Table DP02</td>
</tr>
<tr>
<td>Workers in tract</td>
<td>U.S. Census Bureau, County Business Patterns (Complete ZIP Code Industry Detail File)</td>
</tr>
<tr>
<td>Median earnings (16+)</td>
<td>2017 ACS 5-Year Estimates, Table S2001</td>
</tr>
<tr>
<td>Educational attainment</td>
<td>2017 ACS 5-Year Estimates, Table DP02</td>
</tr>
<tr>
<td>Worked in past 12 months</td>
<td>2017 ACS 5-Year Estimates, Table S2303</td>
</tr>
<tr>
<td>Location of training centers</td>
<td>National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS)</td>
</tr>
<tr>
<td>Household with no vehicle</td>
<td>2017 ACS 5-Year Estimates, Table DP04</td>
</tr>
<tr>
<td>Self-employment</td>
<td>U.S. Census Bureau, Survey of Business Owners and Self-Employed Persons</td>
</tr>
<tr>
<td>Location of childcare facilities</td>
<td>U.S. Census Bureau, County Business Patterns (Complete ZIP Code Industry Detail File)</td>
</tr>
<tr>
<td>Land area</td>
<td>ESRI</td>
</tr>
<tr>
<td>Average family size</td>
<td>2017 ACS 5-Year Estimates, Table DP02</td>
</tr>
<tr>
<td>Percent household children &lt;18</td>
<td>2017 ACS 5-Year Estimates, Table DP02</td>
</tr>
<tr>
<td>Percent household female, with children &lt;18</td>
<td>2017 ACS 5-Year Estimates, Table DP02</td>
</tr>
<tr>
<td>Percent household male, with children &lt;18</td>
<td>2017 ACS 5-Year Estimates, Table DP02</td>
</tr>
<tr>
<td>Childcare costs (U.S. average)</td>
<td>Child Care Aware of America (2017)</td>
</tr>
<tr>
<td>Drug death rate</td>
<td>Centers for Disease Control and Prevention, National Center for Health Statistics, 2018. Compressed Mortality File, 1999-2017 (data file and documentation). Extracted from CDC WONDER Online Database</td>
</tr>
<tr>
<td>Alcohol death rate</td>
<td>Centers for Disease Control and Prevention, National Center for Health Statistics, 2018. Compressed Mortality File, 1999–2017 (data file and documentation). Extracted from CDC WONDER Online Database</td>
</tr>
<tr>
<td>Opioid prescription rate</td>
<td>Federal Reserve Bank of Cleveland (acquired from Centers for Disease Control and Prevention; original source: IQVIA Xponent 2006–2017)</td>
</tr>
<tr>
<td>Renter household</td>
<td>2017 ACS 5-Year Estimates, Table DP04</td>
</tr>
<tr>
<td>Household rent &gt;35 percent of income</td>
<td>2017 ACS 5-Year Estimates, Table DP04</td>
</tr>
<tr>
<td>Household different house same county</td>
<td>2017 ACS 5-Year Estimates, Table DP02</td>
</tr>
<tr>
<td>Household with people &gt; rooms</td>
<td>Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry/ Geospatial Research, Analysis, and Services Program. Social Vulnerability Index</td>
</tr>
</tbody>
</table>
### Table B1 (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homeless counts</td>
<td>U.S. Department of Housing and Urban Development, Office of Policy Development and Research, Enterprise Geospatial Information System, Continuum of Care (CoC) Grantee Areas</td>
</tr>
<tr>
<td>Percent 18–64 with disability</td>
<td>2017 ACS 5-Year Estimates, Table DP02</td>
</tr>
<tr>
<td>Percent 18–64 with ambulatory disability</td>
<td>2017 ACS 5-Year Estimates, Table S1810</td>
</tr>
<tr>
<td>Percent 18–64 with cognitive disability</td>
<td>2017 ACS 5-Year Estimates, Table S1810</td>
</tr>
<tr>
<td>Age-adjusted mortality rate</td>
<td>Centers for Disease Control and Prevention, National Center for Health Statistics, 2018. Compressed Mortality File, 1999-2017 (data file and documentation). Extracted from CDC WONDER Online Database</td>
</tr>
<tr>
<td>Disease prevalence (Table 3)</td>
<td>CDC, 500 Cities Initiative</td>
</tr>
<tr>
<td>Percent household public assistance</td>
<td>2017 ACS 5-Year Estimates, Table B19058</td>
</tr>
<tr>
<td>Percent household TANF</td>
<td>2017 ACS 5-Year Estimates, Table DP03</td>
</tr>
<tr>
<td>Percent household SNAP</td>
<td>2017 ACS 5-Year Estimates, Table DP03</td>
</tr>
<tr>
<td>Percent household SSDI</td>
<td>2017 ACS 5-Year Estimates, Table DP03</td>
</tr>
</tbody>
</table>
Appendix C
Variable Construction

The survey comments and analyzed text associated with those comments are qualitative data. However, a meaningful analysis of differences between LMI and non-LMI tracts requires a quantitative comparison. Therefore, I identify quantitative data that reflect the sentiments expressed in the qualitative comments. In some cases, these quantitative measures are easy to identify and use. For example, the census asks directly about the presence of a disability and whether or not households have access to a vehicle. In other cases, quantitative proxies for the qualitative data are not readily available and must be constructed. I construct several of the proxies used in the text by constructing new data from existing data. This appendix provides details about the construction of the quantitative proxies for qualitative responses in these cases.

Workers per resident in a tract

I construct the number of workers per resident in a tract from Zip-code-level data in the U.S. Census Bureau’s County Business Patterns (CBP) and the ACS. I determine the number of people who work in a Zip code from CBP and divide that number by the resident population in the Zip code from the ACS. The result is the number of people who work in the tract per person living in the tract. I then overlay a census tract layer on a Zip code layer in geographic information systems (GIS) software and assign to the tract the average value of workers per resident for the Zip codes in which it intersects.

Minimum distance to training facility

Information about education and training facilities is extracted from the Integrated Postsecondary Education Data System (IPEDS). IPEDS includes several tables of institutional characteristics. Among these are the programs offered and the geographic coordinates of each education and training institution in the United States. Using the geographic coordinates, I create a GIS layer with the physical location of the institutions meeting my criteria—specifically, institutions that offer no degree higher than an associate’s degree and that offer occupational and basic adult education. To compute the density, I count the number
of facilities in each tract and divide that number by the land area of the tract in square miles.

For proximity, I identify the distance between each tract and the closest institution to that tract. The relevant distance is determined by the minimum perpendicular distance from points representing institutions to any boundary line of the tract. The distance is calculated relative to the boundary of the tract, not the centroid of the tract. This distinction is of little consequence in urban areas where census tracts are quite small in land area, but could be meaningful in rural areas with especially large census tracts. If an institution is within the boundaries of the tract, the tract is assigned a distance of zero.

Self-employment rates

Data useful for calculating self-employment rates are available at the county level from the U.S. Census Bureau’s Survey of Business Owners and Self-Employed Persons. Data are derived from a survey of a random sample of businesses selected from a list of all firms operating during the year with receipts of $1,000 or more, except those classified in a small set of North American Industry Classification System (NAICS) industries. The firms list is compiled largely from IRS data, such as Schedule C filings. I calculate the number of business establishments with no payroll in each county and divide by the working-age population in the county. The resulting self-employment rates are assigned as the self-employment rate for all census tracts in the county.

Density of childcare establishments

The density of childcare facilities is the number of childcare facilities per square mile. These establishments primarily engage in providing daycare of infants or children and are listed under NAICS code 624410. To construct this measure, I collect data on the number of childcare establishments in each Zip code from CBP. I then divide by the land area of the Zip code (in square miles) to get a density. Data on childcare establishments are not available in some Zip codes. In these cases, I compute the density of childcare establishments at the county level and assign that value to Zip codes where Zip-code-level data are not available. Finally, using GIS, I overlay a census tract layer and assign to each tract the average density of the Zip codes in which it intersected.
Childcare costs

The national average reported in the text is “an average of averages”—that is, the average of the average cost of childcare for infants, toddlers, and four-year-olds in center-based and family child care homes. The data come from surveys of state Child Care Resource and Referral agencies reported in Child Care Aware of America (2017). States were asked to provide 2015 cost data for infants, toddlers, four-year-old children, and school-age children for licensed programs or child care programs that are legally exempt from licensing. I use the 2015 average for consistency with the 2017 ACS data, which cover 2013–2017.

Tract crime rates

Crime statistics are not routinely collected at the tract level. I build a predictive model using tract-level crime rates from the 2000 National Neighborhood Crime Study (NNCS) conducted by Peterson and Krivo (2000). I estimate the model using year 2000 data. I then estimate tract crime rates for 2008–12 by employing data from that period in the estimated model. I use 2008–12 data (commensurate with the 2012 ACS) because the goal is to proxy for people with criminal convictions who are free to seek work. These individuals presumably would have already endured the consequences of their crimes, meaning the crime would have been committed well in the past.

The model I construct follows the general logic in the NNCS study. The variables used to predict crime rates are represented by $Z$. The model for estimating year 2000 crime rates is:

$$\text{CR}_{2000,i} = \Phi'Z_{2000,i} + u_i,$$

where $\text{CR}_{2000,i}$ is the crime rate in tract $i$ in 2000, which comes directly from the NNCS data set, $Z_{2000,i}$ represents the factors expected to be predictive of the crime rates in 2000, and $\Phi$ is the set of coefficients for the factors in $Z_{2000,i}$ that I estimate. Because the purpose of the analysis is to obtain crime rate estimates, not to uncover the determinants of crime rates, I do not discuss the results in this appendix. However, the variable list, sample statistics, and model estimates are available from the author upon request. Estimated crime rates for 2008–12 are calculated using the estimates from equation (C-1):

$$\tilde{\text{CR}}_{2008-12,j} = \tilde{\Phi}'Z_{2008-12,j}$$

where the tilde represents a bootstrap estimate.
Rates and densities of homelessness

The only available, consistent source of data on homelessness are counts of people experiencing homelessness that occur as part of the HUD’s CoC Program. Using GIS and CoC region boundary files that include point-in-time homeless counts, I calculate the density of homeless people per square mile (homeless density) and the number of homeless people per 100 residents (homeless rate) for each CoC region, which can be quite small in some cities but quite large in many outlying areas. The counts include both sheltered and unsheltered homeless. I then overlay census tract boundary files and assign the homeless density and homeless rate in the CoC region in which the census tract is located to that tract. Similar values are calculated for the chronically homeless. A “chronically homeless” person has a disability and has lived in a “shelter, safe haven, or place not meant for human habitation” for 12 continuous months or on four separate occasions in the previous three years that total at least 12 months (“Continuum of Care,” 24 CFR 578 Revised July 31 2012).
Endnotes

1 Census tracts are similar to communities in that they were designed to be as homogeneous as possible in their sociodemographic characteristics.

2 The working-age population in most official U.S. labor market statistics is the civilian population age 16 and older. Most people age 16–17 are full-time students, while most people 65 and older are retired. In early 2019, only 22.5 percent of people age 16–17 and only 20 percent of people 65 and older were either working or looking for work.

3 Dynamically, unemployment and labor force nonparticipation are not entirely separable. Individuals with criminal convictions may become so discouraged by their inability to find a job that they quit looking altogether. In that case, they are “discouraged workers” who would be considered “marginally attached” to the labor force but classified for statistical purposes as “not in the labor force.” To be officially classified as unemployed, individuals must have looked for work—specifically, they must have filed a job application—in the past four weeks.

4 This statistic should not be confused with the unemployment rate, which is the share of those in the labor force (employed or officially unemployed) who are officially unemployed. In November 2019, the U.S. unemployment rate was 3.5 percent.

5 I use “mental” as a single term in the text analysis to better distinguish between mental health and physical health concerns.

6 Certain types of industries are more likely to offer lower-skill jobs that LMI workers, who generally have lower job qualifications, can attain. As used here, a “low-skill” job is one that does not require a formal credential or specific experience. These jobs are relatively more common in the health-care, retail, and accommodations and food services industries. Even within industries more likely to hire LMI workers, jobs seem to be more widely available in non-LMI tracts.

7 As a rough estimate of the variation in wages for jobs in LMI tracts compared with jobs in non-LMI tracts, I divide the total payroll of business establishments in each county by the number of establishments with paid employees in the county. Employer-based wage data are available only at the county level, and for this exercise I consider the labor market to be the county in which a tract is located. Under this accounting, the average wage is $41.43 in LMI tracts and $42.10 in non-LMI tracts. My interpretation is that LMI tracts are only modestly more likely to be in low-wage counties than are non-LMI tracts. The data for this calculation are from the Survey of Business Owners and Self-Employed Persons (U.S. Census Bureau).

8 The latest available data (2015) indicate that 50.2 percent of custodial parents (the parent living in the household) have either legal or informal child support agreements with the noncustodial parent (Grall 2018). Just over 80 percent of custodial parents are mothers, 52.7 percent of whom have child support agreements. Among custodial parents with child support agreements, 69.3 percent
receive some payments for child support, but only 43.5 percent receive full child support payments.

9The focus groups were held in March and April 2019 in Chicago, Denver, Detroit, and Kansas City. Transcripts are currently being analyzed; we are not yet able to draw conclusions from the data.

10Due to registries, an exception is sex offenders, but the registries are maintained by individual counties and collecting this data would be intractable. Moreover, most crimes are not sex crimes, and little direct evidence suggests a correlation between the location of sex offenders and non-sexually motivated crimes.

11Renter-occupants are more mobile than owner-occupants. Moving to a different house in the same county proxies for reluctant moves. Desmond and Gershenson (2016) find the likelihood of workers who experienced a forced move losing their jobs to be between 11 and 22 percentage points higher than for comparable workers who did not.

12Age adjustment eliminates the effects of age from crude mortality rates to allow for meaningful comparisons across populations with different underlying age structures. For example, comparing the crude rate of heart disease in Florida to that of most other states would be misleading because of the relatively older population in Florida.
References


**STATEMENT OF OWNERSHIP**

Title of publication: Economic Review  
Publication no.: 417-850  
Date of filing: October 1, 2019  
Frequency of issue: Quarterly  
No. of issues published annually: 4  
Annual subscription price: Free  
Mailing address of known office of publication: Federal Reserve Bank of Kansas City, 1 Memorial Drive, Kansas City, Missouri 64198-0001  
Mailing address of headquarters or general business office of publisher: Federal Reserve Bank of Kansas City, 1 Memorial Drive, Kansas City, Missouri 64198-0001  
Publisher: Federal Reserve Bank of Kansas City, 1 Memorial Drive, Kansas City, Missouri 64198-0001  
Editor: Elizabeth Cook Willoughby, Federal Reserve Bank of Kansas City, 1 Memorial Drive, Kansas City, Missouri 64198-0001  
Managing editor: Jonathan L. Willis, Vice President, Federal Reserve Bank of Kansas City, 1 Memorial Drive, Kansas City, Missouri 64198-0001  
Owner: Federal Reserve Bank of Kansas City, 1 Memorial Drive, Kansas City, Missouri 64198-0001  
Issue date for circulation data below: Third Quarter 2019

<table>
<thead>
<tr>
<th>Extent and nature of circulation</th>
<th>Average no. copies of each issue during preceding 12 months</th>
<th>Actual no. copies of single issue published nearest to filing date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of copies</td>
<td>1,800</td>
<td>1,800</td>
</tr>
<tr>
<td>Paid and/or requested circulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paid or requested mail subscriptions</td>
<td>1,146</td>
<td>1,101</td>
</tr>
<tr>
<td>Total paid and/or requested circulation</td>
<td>1,146</td>
<td>1,101</td>
</tr>
<tr>
<td>Free distribution by mail (samples, complimentary and other free)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Free distribution outside the mail (carriers or other means)</td>
<td>165</td>
<td>171</td>
</tr>
<tr>
<td>Total free distribution</td>
<td>165</td>
<td>171</td>
</tr>
<tr>
<td>Total distribution</td>
<td>1,311</td>
<td>1,272</td>
</tr>
<tr>
<td>Copies not distributed</td>
<td>489</td>
<td>528</td>
</tr>
<tr>
<td>Office use, leftovers, spoiled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1,800</td>
<td>1,800</td>
</tr>
<tr>
<td>Percent paid and/or requested circulation</td>
<td>87.4</td>
<td>86.5</td>
</tr>
</tbody>
</table>
Spending Patterns and Cost of Living for Younger versus Older Households

Payment Card Fraud Rates in the United States

Why Aren't More People Working in Low- and Moderate-Income Areas?