TRACKING THE WORK HISTORIES OF LOW-WAGE EARNERS

An Undergraduate Research Scholars Thesis

by

ROGER "KIRK" REESE, JR

Submitted to Honors and Undergraduate Research
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by
Research Advisor: Dr. Jonathan Meer

May 2015

Major: Economics

TABLE OF CONTENTS

ABSTRACT

Tracking Work Histories of Low-Wage Workers. (May 2015)

Roger "Kirk" Reese, Jr
Department of Economics
Texas A&M University

Research Advisor: Dr. Jonathan Meer
Department of Economics

It is critical that policy makers and researchers have an idea of what future earnings look like for a typical low-wage worker, and how this picture differs for individuals of various socio-demographic characteristics. In light of the growing income gap, scholars have done well to measure growth in overall income mobility. These measures have been improved by lifetime, longitudinal data from IRS tax records, but these administrative data offer little demographic detail. The panel data that do offer such demographics, lack in length coverage or sample density. We find either short temporal coverage of study, or what could be characterized as a time-series of cross-sections, drawing conclusions from one individual's earlier career and another's later career. Furthermore, the relative and overall mobility approach that has been the focus of existing literature, tells us nearly nothing about the absolute experience of low-wage individuals hoping to climb the ladder to prosperity. Also, little has been done to de-aggregate average earnings across specific sub-sets of the low-wage population. The Survey of Income and Program Participation panel offers a balance in the trade-off between years of coverage and demographic information. The SIPP Synthetic Beta preserves many detailed variables of the SIPP, while linking observations to lifetime administrative data from the IRS and Social Security Administration. These detailed demographics allow the separation of typical low wage workers

from high-wage/high-skill workers temporarily receiving low wage. I use the SSB to map out how low-wage workers, conditioned as being so by their wage rate and work consistency, can expect their absolute earnings to progress throughout the remainder of their career. This paper does not attempt to propagate a particular point, but rather to take an explorative approach to usefully organizing and making available insight into long-term earnings trajectories for different types of low-wage earners.

SECTION I

INTRODUCTION

The promise of prosperity for consistent hard work is at the heart of this American economy. It is the promise that keeps low-wage workers working, in a nation that would provide them similar income for taking long, leisurely breaks from employment. It is also the promise that makes the conditions incurred by the lowest earners seem more tolerable on the conscious of society. But, how good a promise is it? From childhood to adulthood, I observed as each of my parents experienced what we call the American dream, rising from meager combined income that barely rented a two bedroom house and fed three children, to each individual having an income that would sustain a single person in the lower cohorts of the upper class. This common motif can be deceptive in regards to the real experience for a low-wage worker, the same way predictions from earnings data can be deceptive in regards to the expected outcome for a low wage worker.

The catch to my parents' story is that, at their time of being low-wage workers, they were actually high wage earners, temporarily preoccupied. You see, both were busy earning a higher education beyond a bachelor's degree. Their story is not indicative of what we would expect for a worker earning wages expected by their skill level. That is the story that I am interested in. What does a typical low wage worker at the present, expect to earn in the future. Additionally, how is this expectation different for different "sorts" of people. For example, by everything we have learned about inequality, we would not imagine that a black female earning a low wage at the present, would expect to see the same gains over the next ten years as a white male having the same wages and type of job. The results may be surprising.

This promise of increasing wages, the American Dream as it is referred to, is a hot topic in economics. The topic is typically discussed, however, focusing on overall mobility. Overall, relative mobility to be even more specific. While this device offers much insight into the inequality, particularly how the growing disparity between the nation's richest and poorest is felt –or not noticed- overall, it provides very little insight into the actual experiences of individual workers. It is interesting to know how rapid upward movement is throughout the economy, but less so to someone focused on individuals' outcomes. By measuring income mobility overall, we can easily miss how one income level differs from another. By measuring income mobility relatively, we can't say explicitly how well someone is likely to do, only how likely they are to do better than their neighbor. The usefulness in regards to policy that would affect low-income workers is limited by other information, such as what level new entrants in the workforce are starting at, and how many are entering. Looking at absolute earnings trajectories creates a clear and simple picture for policy makers wanting to protect the individuals most sensitive to bad policy resulting from misunderstandings over the expected earnings levels.

Surely, using historical earnings to predict what someone earning in a particular level now, is likely to make a number of years into the future has been done for as long as economics and earnings data have coexisted. This approach has several key advantages. Primarily, the data set used allows for low-wage workers to be conditioned on wages, while still projecting earnings fifteen years into the future. Of course we are conditioning low-wage workers on wages, right? But traditionally (with exception), researchers must choose between short-term panel datasets with wages or hours worked and other work descriptive variables, or the long-term coverage of

administrative data that only provides earnings. The issue with this is that the individuals you'd select from the sample, because their earnings were low enough, would include high skilled, high paid workers that worked few hours. You would also get people like my parents (and myself): higher-earners temporarily experiencing low wages. If you wanted to combine the long term earnings information from one data-set with the wage and work hour information from another, you would be doing so across individuals. The unique data-set known as the SIPP-Synthetic Beta allows us to tie the long-term future earnings records for an individual to the wages and hours for that same individual observed in the extensive Survey of Income and Program Participation panel.

The few datasets that offer this unique combination are quite thin in sample size compared to the SIPP derived SSB, becoming unreliable when our attention is directed towards very specific types of individuals. They also provide a plethora of other problems that will be discussed later. A secondary objective of this project is to lay the ground work for future research that would target the outcomes of these specific and interesting "types" of people, to see how their prospect for earnings advancement differs from low wage workers as a whole. In this paper, only gender and race will be demonstrated.

Existing mobility literature will benefit from the addition of absolute earnings projections for individuals that are presently typical low-wage workers, as confirmed by their wages and work hours at time of observation, to better understand what kind of absolute upward mobility these individuals can reasonably expect to obtain in years after which they were observed as such. In this paper, I will first conceptually explain income mobility and the existing state of the

literature. After reviewing the literature, the strategic approach to this project will be defined, and the key piece of this project, the dataset, will be detailed in my methods section. In the Analysis section you will find the specific techniques applied to the data, and see the final trajectories outputted for the low wage workers and some interesting results for certain sub-sets. Finally, findings will be interpreted and discussed in the concluding section while and criticisms and cautions will be expanded upon.

SECTION II

LITERATURE REVIEW

This paper is quite different from the contemporary work in income mobility, in that it doesn't directly look at a measurement of mobility, but rather seeks to see how this mobility is realized through growth in income. Understanding the existing literature though creates a context for the results of this paper. It is be best to begin by explaining what income mobility is.

The economy moves everyone with it, pushing participants up as it expands relative to the population, and pulling them back down as it recedes. The American Dream implies however that an individual can move forward faster than the economic current carries them in good times, or otherwise continue to rise -or reduce the rate they fall- in times of economic contraction, by adjusting their efforts and working harder or more efficiently. The combined movement of the individual from both effects is what we call "income mobility." Auten & Gee (2009) utilize an excellent metaphor of income mobility as people on an escalator. The escalator's travel represents the movement due to economic conditions, but people can move forward or backwards relative to the escalator by stepping up or down with a range of speeds.

Income mobility is primarily approached in two ways, relative and absolute. The relative measure is a rank-based construct. In this relative measure, people are ordered according to income, and then separated into quantiles, and the probability of moving from one quantile to higher one in a certain time frame is utilized as a measure of mobility. Movement to even higher groups is given more weight to reflect a greater amount of mobility. This measure is a favorite in

mobility research. We will be focused on absolute mobility, though. This measure, as the name implies, is not relative, but looks at how far an individual has traversed in actual dollars earned. We can also draw a distinction between short-term and long-term mobility. The short-term approach attempts to measure mobility within a year or across only a few years, whereas long-term mobility is measured across time periods closer to the working span of a person's life. It seems to me that short-term mobility (up or down) would be most useful as a measure of wage volatility or instability, whereas long-term measurements are a better view of mobility as it is considered here.

The growing disparity between low and high income groups, commonly referred to as the income gap - see Katz & Autor (1999) - poses a real threat to this mobility; it implies that the step an individual must take to get from one state of financial wellbeing to the next is becoming larger. If this step is getting larger, one could intuit that individuals moving upward at the same rate would not get as far over the same amount of time. If they are moving faster as well though, then the growth in the income gap would be offset. This is where the debate has been centered. Which claim is better supported by the research? The existing literature predominantly agrees that mobility overall is still alive, and in fact, relatively unchanged, as confirmed by Kopczuk, Saez, and Song (2010) using longitudinal Social Security Administration earnings records to measure long term, rank-based (relative) mobility. As it turns out, the steps that individuals are actually taking on average are becoming larger as well. So, the distance they travel across the income distribution, proportional to the size of the distribution itself, has remained largely unchanged. They also confirmed the prior results (Gottschalk, 1997), that short-term, relative mobility has been constant since the '70s. Indeed this is remarkable knowledge to have, but what

I found even more interesting about their work, was the decreasing gender gap in mobility. While the overall mobility is constant, female workers had an unprecedented increase in mobility ratings, while men experienced a decrease of similar proportions.

Some have shown a similar interest in mobility, but with more focus on the lower parts of the income distribution. Auten & Gee (2013) confirm the findings of Kopczuk, Saez, and Song, that long-term mobility increased to offset the income gap, but with income-tax data rather than SSA data. They further conclude that starting position is one of the most significant factors in determining mobility, and that there is considerable upward movement out of the bottom; Real incomes increased for all tax-payers the period of study, about half of those in the bottom quintile in 1996 moved to a higher income within roughly a decade. While this sheds light on the amount of mobility out of low-wage jobs, and how overall mobility is spread across income levels, we still lack actual numbers as to what these individuals ended up earning.

Colin Campbell (2011) takes a solid swing at unveiling how low wage people of different educational levels, race, and areas compare in "Low-Wage Mobility during the Early Career." While he uses a rather thin version of the already more lightly sampled PSID panel, having eliminated all individuals from the sample who are not white or black, he still reaches some rather interesting results about the remaining subset. White Americans can expect considerably higher income mobility than Black Americans. Accord to his work, 76% of those starting under $12 per hour made it to $12 or higher over the 9-year period. Again, initial position is one of the largest factors in determining mobility, and also upward movement out of the low-income range

is less permanent for black Americans, women, and the less educated. It will be important later to mind the results of Colin's paper.

SECTION III

METHODS AND DATA

There is a well-known trade-off between longitudinal administrative data and panel data. Panel data typically offers a rich range of detailed variables, while covering rather few years. Typical longitudinal data on the other hand can offer highly accurate information across large spans of time. I will combine the advantages of long-term administrative data with demographic variables available in shorter term panel data to construct absolute wage trajectories that will allow us insight into what a career wide mobility actually looks like for the typical low-wage worker. The methods discussed will allow us to condition our subset to be a much more accurate representation of low wage workers, and to take a much more comprehensive look at how the upward mobility experienced by initially low-wage earners de-aggregates across many sets of demographic characteristics.

Individual tax and social security records have been a popular administrative data set for income mobility research, particularly for those interested in long-term mobility. The lack of detail and demographic information provided in these records is underwhelming though. On the other hand, there are panels such as the Survey of Income and Program Participation which offer more information than any one field of research could possibly digest alone, but cover only a few years of each individual's life. The census bureau has developed a data set called the SIPP Synthetic Beta. The true "gold standard" version of the data ties SIPP participants and all their panel responses to a large number of administrative records from the IRS and SSA. While the public cannot access the full version for security reasons, this "semi-synthesized" version

preserves a large, and growing number of the interesting SIPP variables, while linking the participants to the administrative variables most requested by researchers. Any programs written for use on the SSB can be given to the Census Bureau to be run on the gold standard and they will return the results. The SSB seemed to be just the right balance for this purpose; long-term earnings data, with a much more comprehensive set of demographic variables and detailed work history than have been available in prior studies of long term mobility. Also, the SIPP is designed to oversample low income individuals, making it more accurate for the people we are particularly interested in.

This is an appropriate point to note that the Panel Survey of Income Dynamics, commonly used in research on earnings and mobility, provides wage and hour variables alongside long term earnings data, the distinguishing quality that has been boasted for the SSB. These data differ though the sheer sample size, as well as the consistency over time of the SIPP survey. The PSID makes such drastic changes after 1997, that the before and after data are typically treated as entirely different data sets. Also, the sample is relative thin, and loses value quickly as you condition the survey's subset on more variables.

I want to know, if we observe a low wage worker conditioned on having particular characteristics at a point in time, what they are earning five, ten, to fifteen years later. The methodology is to use the variables observed during the SIPP to "type" a person, then see how their earnings progressed following their initial observation as this type of low-wage worker. For example, what can a white, female worker who is employed in a non-managerial and non-professional position expect to be making ten years into the future? The project will involve only

individuals who were low wage-workers during their SIPP panel. Furthermore, we will remove those individuals who are not actually typical low-wage workers, but whose earnings or wages might classify them as such in a less detailed dataset. This would include people like college students, who might be experiencing low-wages, despite being a high wage worker under normal conditions. Such individuals would raise the future earnings predicted for a low wage earner, uncharacteristically.

The extensive demographic variables available in the SIPP are only available during the panel's coverage of that individual. Therefore, we cannot see when they cease to have this status later in life, or if an individual falls into another "type" beyond their observation in the SIPP. This limitation supports the time-since-observation approach, as opposed to years of work experience. Another key consideration is that while we can observe a wage – or rather a wage equivalent, as will be explained- at a monthly frequency during the SIPP, after an individual leaves the SIPP, we are limited to annual earnings and quarters worked. So, our primary interest is on future earnings, but the detailed information in the SIPP allows the construction of strong criteria for what it means to be observed as a typical low-wage worker at time zero.

By establishing criteria that someone was a particular type of person during their observation in the SIPP, and looking at how the average earnings for that person change over time after being leaving the SIPP, the problem of people changing types after their participation in the panel just becomes noise; it is essentially part of the probability of their having a particular earnings level at a certain time. This may not be ideal, but it still allows us to see things that would have otherwise remained unobserved. Later projects could use econometric techniques to impute other

information, which allows us to take into account the differences in probability of having these changes across groups at certain periods of time. For this paper the assumption is made that the probability of changing status remains fairly constant across "types."

It even seems reasonable to extend our criteria from those who were in school or otherwise non-typical low wagers, to even removing all part-time and other irregular workers as well, because we are looking at earnings rather than wages for the time after panel participation, and someone who has established a reputation as a consistent part time worker during the SIPP is more likely to continue being a part time worker in the future, their earnings will distort the average future earnings of a typical low wage worker. Similarly, people with irregular hours throughout the SIPP are likely to have irregular earnings in the future. For example, someone who made ten or twelve dollars per hour, but only works 10 hours a month, is likely otherwise supported, and could be doing some odd jobs for supplemental income. Again, this person doesn't characterize a typical low wage worker, so we would not want them included in consideration of what a low-wage worker can expect to be earning in a few more years of solid work.

As for the conditions of types of workers, the strategy here is to be as modular as possible, so that this project can lay the groundwork for future income and mobility projects. The code is written in such a way that conditions can be switched on and off easily, in order to quickly get different perspective for different types of people, and new variables can be easily added. Because the good people at the Census Bureau are constantly adding new variables to the SBB, this project should continue returns for future research. For right now we have focused only on race, gender, and type of worker (professional/managerial or not), because the primary focus of

this paper was being able to generate accurate conditions for low wage workers. What has been done with the typing of workers though, provides proof of concept, and inspiration for the future.

A few more things to note; The focus was on young workers between the age of 20 and 35, because this would give the furthest look into the future for understanding where a low wage worker will go. Also, averaging annual earnings across different calendar years, and consequently, different states of the economy and places in the business cycle, makes our prediction more accurate in the long run, but at the cost of short run precision. If we experience an era of economic expansion greater than the average over this time period, our prediction would be low, and vice-versa for recession.

SECTION IV

ANALYSIS

Specifically, the process undertaken is as follows: First, because all values are reported in the SIPP and on administrative records as nominal amounts, adjustments had to be made for inflation. These were done using the CPI to adjust for changes in purchasing power. After establishing all earnings and wage measurements in 2013 dollars, the next step was to establish the conditions for being a low-wage worker at the time of observation. This part of the process involves many arbitrary decisions, which I will now explain.

For establishing that an individual qualifies as earning a low-wage at time of observation, their self-reported earnings every month in the SIPP panel are divided by reported hours worked for that month to create a wage equivalent. Individuals whose wage equivalent was over $12 per hour on average over the time we observed them in the SIPP, are flagged for exclusion. This wage was decided on so that these results could be considered relative to those in Collins (2011).

Then, the criteria for being a typical worker is established by flagging for exclusion, anyone that meets one of two criteria. The first criteria is intended to catch anyone that is an irregular worker. It seems reasonable that someone who has a regular job and goes on vacation and gets sick in the same month- the two most normal reasons to have temporarily low work hours, while still remaining regularly employed- could have less than 40 hours worked that month. Two months like this in the same year would be extremely unlikely. Anyone having more than two months with less than 40 hours gets a flag indicating that was a bad work year. The second criteria is to

catch anyone that was consistently part time when observed during the SIPP. For this flag, an individual must have at least six months in the year with less than eighty hours worked. If an individual has the part-time flag for all their years, or the irregular work flag for at least half of their years in the panel, then they would be flagged for exclusion.

The next condition applied is an age filter, to remove anyone over the age of 35. Again the individual is flagged, only if they were over this age for the entire time of the SIPP. All monthly or yearly observations during which an individual was under the age of 20, are excluded from all calculations and weightings, as if they did not exist. Another condition; those who were in school at the time they are observed in the panel are flagged for exclusion. At this point, any number of available demographic conditions can be applied. We limit ourselves to low skill (not a professional or managerial position), gender, and race for the sake of time, but future projects that use this as a base, can easily adapt any other existing variables in the SSB, or any that the census adds in the future at this point in the process.

A variable is created to represent years since leaving the SIPP panel, and annual earnings are averaged across all individuals to this variable. Also, to further improve the accuracy of our predictions, those years for which an individual did not have 4 quarters of employment are not included in the average for that year. All annual earnings are excluding those with zero income, because we only want to know how much people who are actually working consistently expect to earn. Instead, we include the probability of having zero income at that time, as an indication of the proportion of individuals that cease earning after this amount of time. Since wages reported are cut off at the taxable maximum, the probability of reaching this maximum is also included.

The results for all young, low-wage workers that were not in college or in a managerial or professional position at the time they were observed in the SIPP are shown in Figure 1 below.

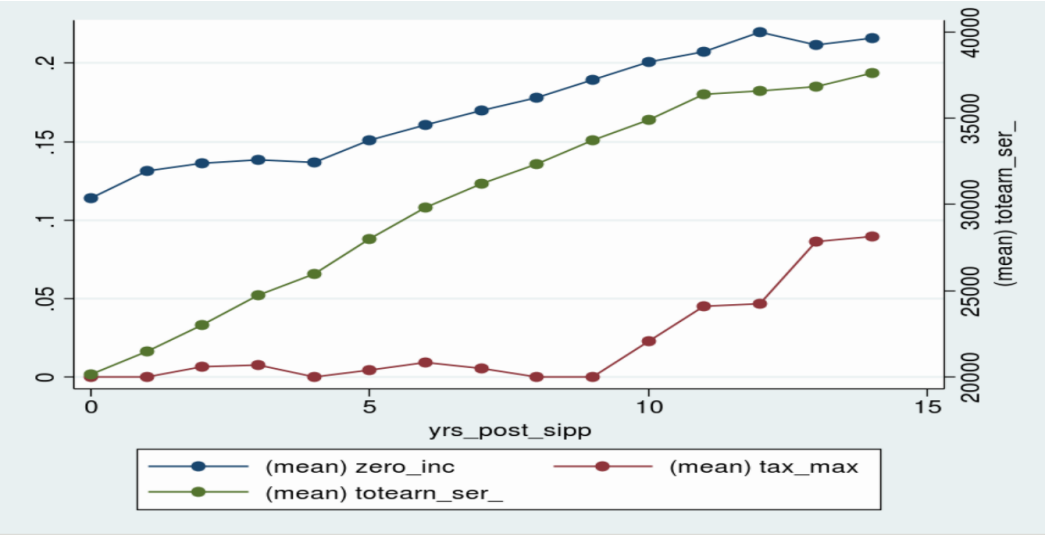

Figure 1: Expected Outcomes for Low Wage, Low Skill Workers, Not in College at Time of SIPP Observation

The middle line is our primary interest, the expected annual earnings. The top and bottom lines are the probably of having zero income and the probability of hitting the maximum taxable income respectively. To further demonstrate the usefulness of this work, I have included the same graph below, but with additional conditions, first you will see the trajectories for all white male workers included in the above sub-sample, then black males, and finally white females, in Figure 2, Figure 3, and Figure 4 respectively.

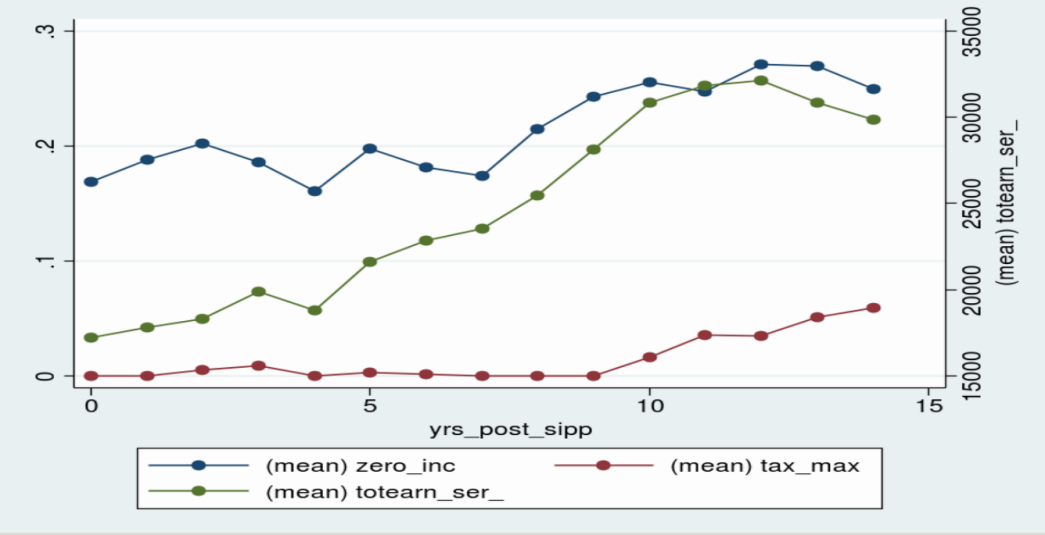

Figure 2: Expected Outcomes for Low Wage, Low Skill, White, Male Workers, Not in College at Time of SIPP Observation

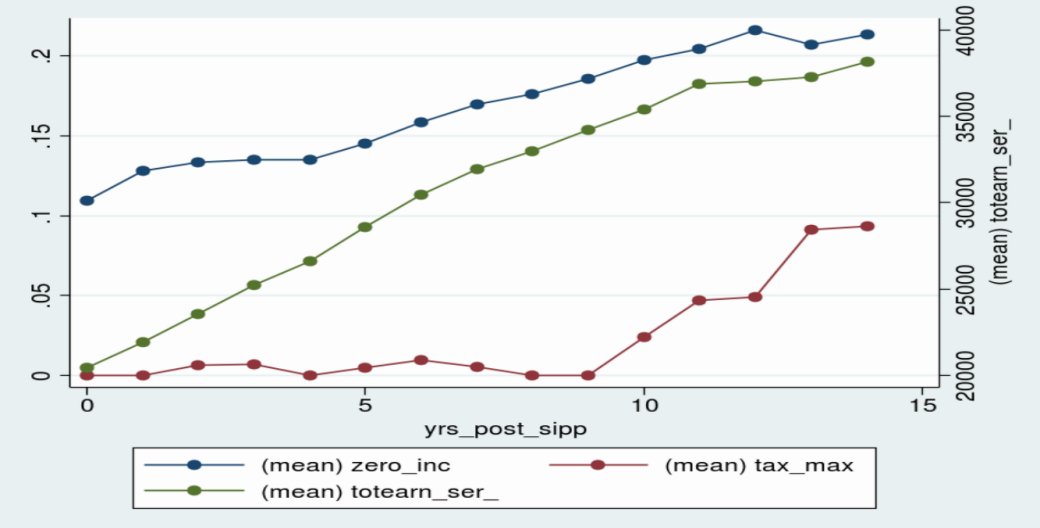

Figure 3: Expected Outcomes for Low Wage, Low Skill, Black, Male, Workers, Not in College at Time of SIPP Observation

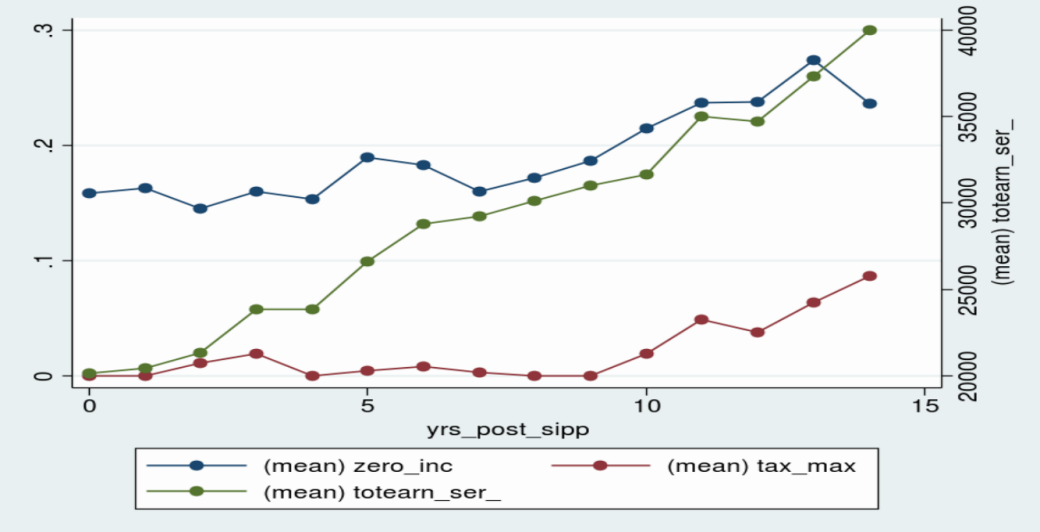

Figure 4: Expected Outcomes for Low Wage, Low Skill, White, Female Workers, Not in College at Time of SIPP Observation

In the appendix you can find the same breakdown for every remaining combination of black, white, hispanic, and male or female, subject to the same low wage, low skill, regular work conditions above.

SECTION V

CONCLUSION

While the results for the general low-wage working population are not particularly surprising, we now have specific numbers for accessing future outcomes for low wage workers. This will be quite useful knowledge for policy makers, think-tanks, future economic research, and planning for the retirement of low wage workers. The linear trend from time of being a low wage worker, until about twelve years later lets us know what has been historically typical for low-wage earners, and will be a benchmark for comparison when looking at rate changes within certain time periods, or phases of the business cycle. With the change of a variable, the program could show us where an older low wage worker came from, so that we might postulate why they ended their career earning low wages, and also to look at middle age low wage workers to assess simultaneously where they came from and where they are going. Also, age conditions could be tightened to remove more retirees, but this would affect only the probability of zero, rather than average earnings.

It is very surprising how badly white males seem to do compared to females and black males. It seems to go against everything that the existing literature would have us believe, directly contradicting the findings of Cambell (2011), that blacks experience a much lesser degree of mobility than whites. Could it be that low wage, white male workers are considerably disadvantaged compared to their female, black, and hispanic counterparts? Or, are the white males who are comparable to the higher achieving female, black, and hispanic workers making higher wages from the start and missing the wage condition entirely, leaving only a lower

achieving subset of the white, male population to compete in the conditional sample. These results ask more questions then they answer, which is convenient since there weren't many questions asked.

Finally, there is a large and obvious issue. Since all variables are averaged across different years, the different states of the economy are not accounted for. If we had a heavy proportion of our sample being 5 years out from their initial SIPP observation as a low wage worker right after the financial crises, then this process would predict a lower expected earning than a low wage worker today might expect in five years. This problem is perpetuated by a lack of symmetry in the business cycle. Using data on the business cycles, or separating our sample into groups observed initially on specific years are ways to deals with this, but weren't feasible for the one year time span allotted for this project. Certainly, these corrections will be implemented in a work that expands on this one, but they should not considerably detract from the results found here, when looking at long-run expectations.

It seems safe to say that we have observed the American Dream for those that have to start at the bottom of the income distribution. There are however some surprises as to who experiences it more strongly. All in all, this project has contributed immediately valuable information to the literature. At least as important as the contribution to our current understanding on mobility for low wage workers was the secondary objective of laying the ground work for future research into income mobility and the income distribution. Many questions have been inspired for future mobility and income research, the results will continue to grow as the SIPP participants age, and

so will the econometric abilities of this author; This will lead to the attainment of many future insights into long-term mobility.

REFERENCES

Auten, G. & Gee, G. (2009). Income mobility in the United States. *National Tax Journal*, 301-328.

Auten, G., Gee, G., & Turner, N. (2013). Income inequality, mobility, and turnover at the top in the US, 1987–2010. *American Economic Review,* 168-172.

Campbell, C. (2011). What is a minimum wage? *Research in Social Stratification and Mobility*, 30(2), 175-185.

Gottschalk, Peter (1997), Inequality, income growth, and mobility: The basic facts, *Journal of Economic Perspectives*, 11 (2), 21–40.

Katz, L., Autor D. (1999). Changes in the wage structure and earnings inequality. In: Ashenfelter O., Card D. *Handbook of Labor Economics*, vol. 3, 1463-1555.

Kopczuk, W., Saez, E., & Song, J. (2010). Earnings inequality and mobility in the United States: evidence from social security data since 1937. *Quarterly Journal of Economics,* 125, 91-128.

U.S. Census Bureau (2015). SIPP Synthetic Beta: version 5.1 [Computer file]. *Cornell University, Synthetic Data Server* [distributor].

APPENDIX

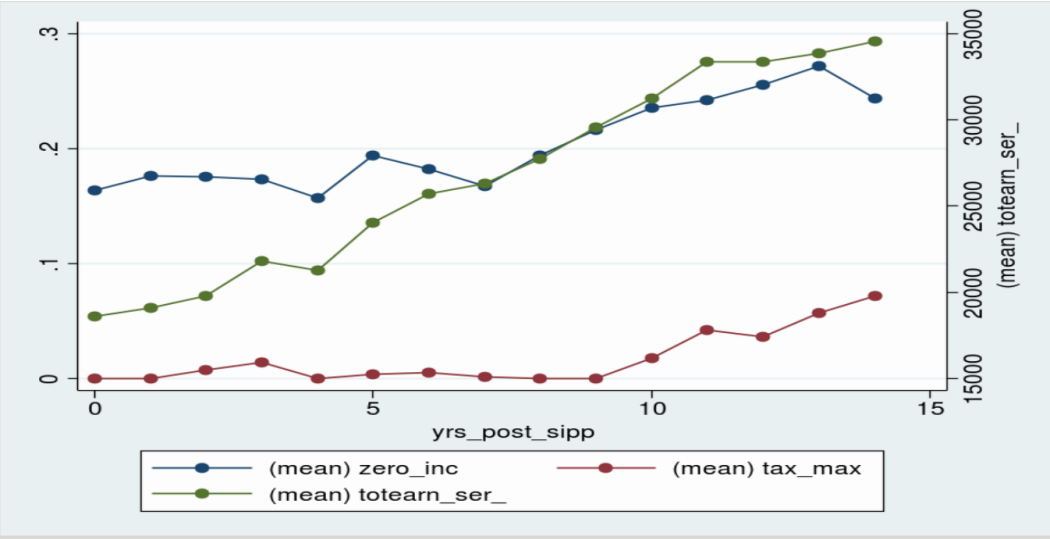

Figure A-1: Expected Outcomes for Low Wage, Low Skill, White Workers, Not in College at Time of SIPP Observation

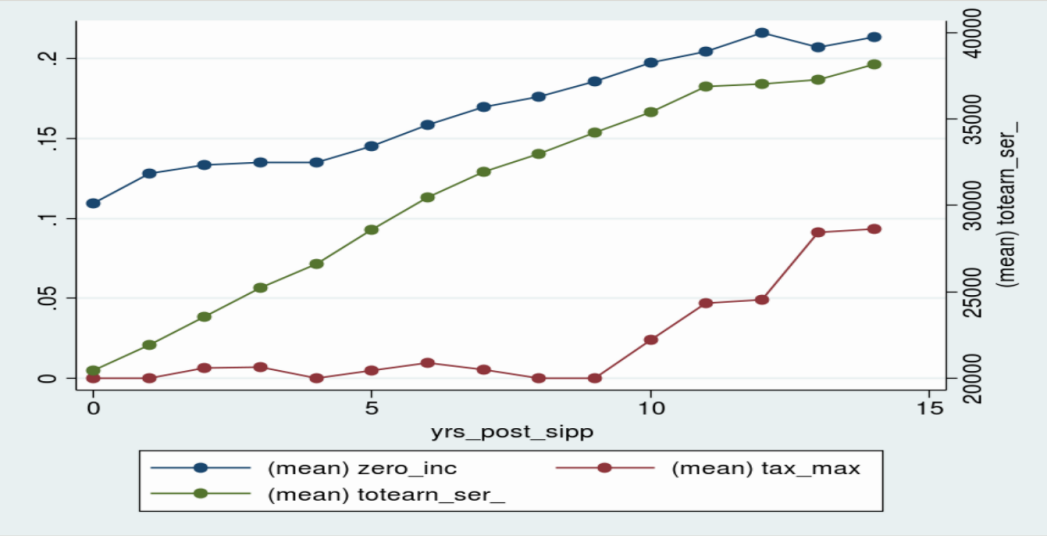

Figure A-2: Expected Outcomes for Low Wage, Low Skill, Black Workers, Not in College at Time of SIPP Observation

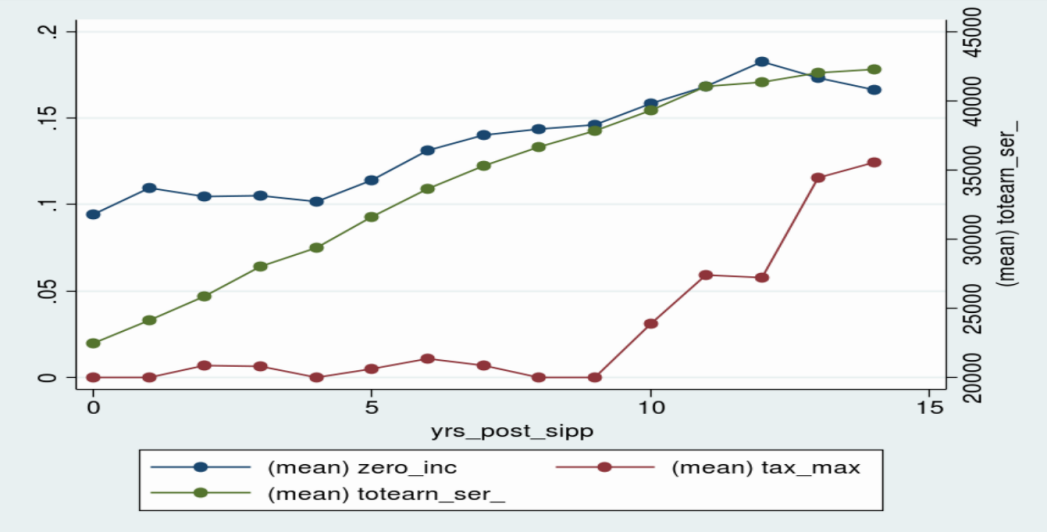

Figure A-3: Expected Outcomes for Low Wage, Low Skill, Female, Black Workers, Not in College at Time of SIPP Observation

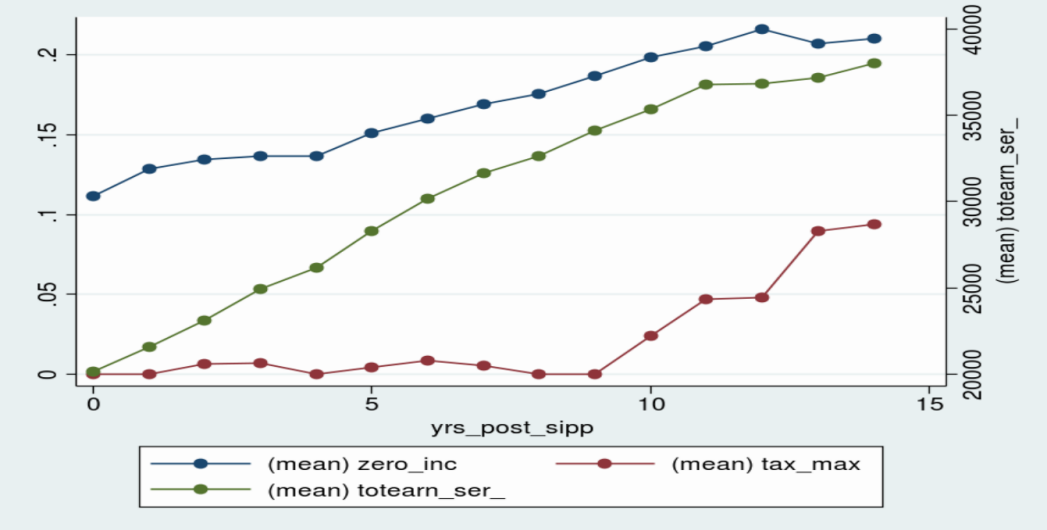

Figure A-4: Expected Outcomes for Low Wage, Low Skill, Hispanic Workers, Not in College at Time of SIPP Observation

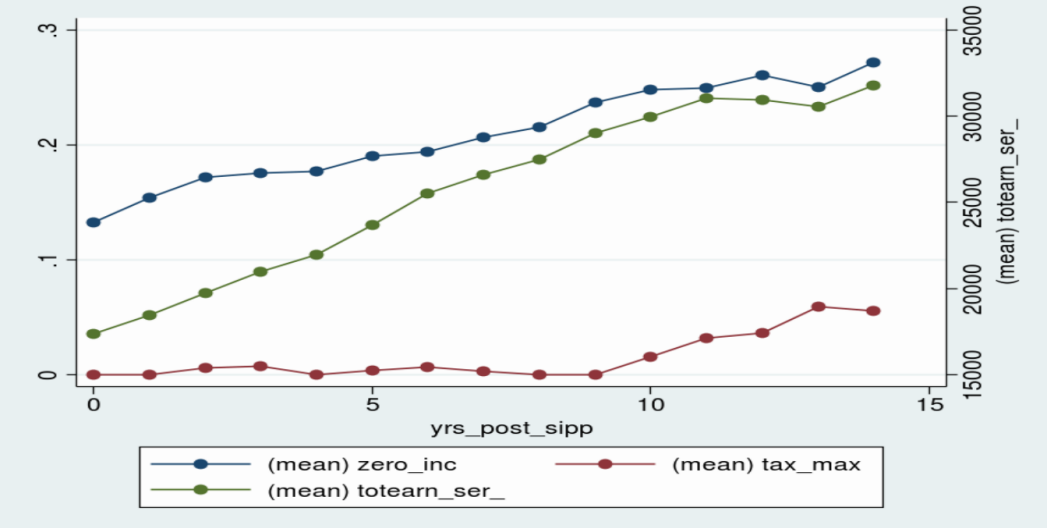

Figure A-5: Expected Outcomes for Low Wage, Low Skill, Male, Hispanic Workers, Not in College at Time of SIPP Observation

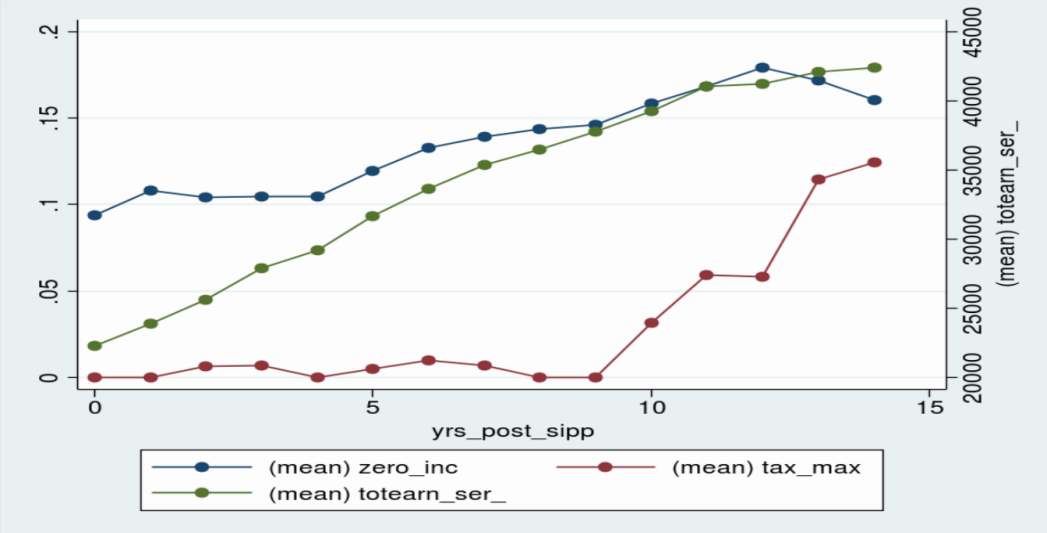

Figure A-6: Expected Outcomes for Low Wage, Low Skill, Female, Hispanic Workers, Not in College at Time of SIPP Observation