

INSTITUTE OF POLICY AND PLANNING SCIENCES

Discussion Paper Series

No. 1039

Negative Self Selection into Self Employment
among African Americans

by

Daiji Kawaguchi

May 2003

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

Negative Self Selection into Self Employment among African Americans

Daiji Kawaguchi¹

Institute of Policy and Planning Sciences, University of Tsukuba
Tennodai 1-1-1, Tsukuba, Ibaraki, 305-8573, Japan

Tel: (81)-29-853-7432

Fax: (81)-29-853-3849

E-Mail:kawaguch@sk.tsukuba.ac.jp

May, 2003

¹I thank David Neumark, Jeff Biddle, Yasushi Ohkusa, and Fumio Ohtake for their comments, as well as the seminar participants at Michigan State University, Osaka University, and the Kansai Labor Forum. I also thank Hisahiro Naito for helpful discussion and Donna Maurer for editorial service.

Abstract

This paper attempts to shed light on the mechanism behind the lower rate of self-employment (SE) among African Americans compared with Caucasian Americans. Both consumer discrimination and discrimination in the credit market, combined with anti-discrimination law enforcement in salary/wage (SW) sector, explain why African American workers, eligible African American workers in particular, are less likely to be SE because the cost of being discriminated against is high. A previous influential study tested negative self-selection into SE among African American workers using Heckman's sample selection correction under certain excluded variable assumptions. Using matched CPS panel data, this paper tests the same prediction without relying on the excluded variable assumptions. More specifically, current salary/wage workers are divided into prospective SE workers and prospective salary/wage workers and the distributions of current earnings of these two groups are compared. The analysis reveals both positive and negative self-selection into SE among Caucasian Americans, but almost only negative self-selection among African Americans. This finding is consistent with the theoretical predictions of consumer and credit market discrimination against African American self-employed workers and confirms the empirical results of the previous study.

1 Introduction

During the 1990s, self-employed (SE) workers persistently consisted of about 12 to 13 percent of the male work force in the US. The breakdown of this number into racial groups indicates a significantly different SE rate for the various racial groups. While 13 to 14 percent of the non-Hispanic Caucasian male work force was SE, the number for African Americans was between 5 and 6 percent (See Fairlie (2000b)). This low rate of SE among African American has attracted much attention from the media and academicians because SE is often considered as a vehicle of upward social mobility for ethnic and racial minority group members that enables them to circumvent market discrimination.¹ In addition, the differential likelihood of small business ownership creates tensions among the various ethnic and racial groups, as documented in Fukuyama (1995).

The low SE rate among African Americans is rather surprising because African American workers can avoid employer discrimination by starting their own businesses. If SE is used as a means to avoid employer discrimination by African American workers in the absence of discrimination in other markets, then the SE rate among African American should be higher.²

Other forms of discrimination against African American SE workers, though, explain the lower SE rate among them. Namely, (i) consumer dis-

¹Evidence for young, less educated workers is provided in Fairlie (2000a)

²This possibility is pointed out by Moore (1983); however, he did not find evidence that African American workers use SE as a means to avoid employer discrimination.

crimination against African American SE workers as suggested by Borjas and Bronars (1989), and (ii) small business credit market discrimination against African American as documented by Blanchflower, Levine, and Zimmerman (1998). Fairlie (1999) revealed that both a lower rate of entry into SE and a higher rate of exit from SE by African American workers contribute to the lower rate of SE among them using PSID. He found that the racial difference in asset holding and the probability of having a self-employed father explains between 15 and 30% of the gap in the entry rate. However, the rest of the difference in the entry rate and almost all the difference in the exit rate between African Americans and Caucasian Americans could not be explained by differences in observable characteristics. He then speculated that the remaining gap would be due to consumer or credit market discrimination. Kim and Deltas (2002) also carefully compared the transition rates into and out of SE by Caucasian Americans and African Americans, using matched CPS files, and they observed a lower entry rate into and a higher exit rate from SE among African Americans, as in Fairlie (1999). Due to the large data set, they could separate the sample by age groups and found evidence of a declining entry rate difference across racial groups as workers age. They speculated from this evidence that African American workers accumulate assets as they age and mitigate liquidity constraints due to credit market discrimination. This paper complements the suggestive evidence by Fairlie (1999) and Kim and Deltas (2002) by directly testing the implications of consumer and credit market discrimination for potential African American SE workers.

As discussed in the following sections, the existence of consumer and credit market discrimination commonly predicts negative self-selection into SE among African American workers. While empirically distinguishing these two separate hypotheses is beyond the scope of this paper, the analysis offers clear evidence that is consistent with consumer or credit market discrimination theories without relying on restrictive assumptions by exploiting the features of panel data.

The rest of the paper is organized as follows. Section 2 introduces theories predicting both a lower SE rate among African Americans and negative self-selection into SE. Section 3 critically reviews the previous evidence on the negative self-selection into SE among African American workers. Section 4 proposes a more robust method of identifying the negative self-selection into SE among African Americans. Section 5 discusses the data used in this study. Section 6 explains the empirical results based on matched CPS files. Section 7 concludes.

2 Negative Self-Selection into SE among African Americans

2.1 Consumer Discrimination

Borjas and Bronars (1989) used the search theoretic framework to describe how the consumer discrimination of majority group members against minority group members discourages eligible minority workers from starting up their own businesses. They model SE workers as agents who produce goods

with heterogenous productivity and sell the products with homogenous sales ability, allocating their limited time to production and selling activities optimally. In the model, majority consumers are agents with discriminatory taste who buy from minority sellers only when the price charged by the minority seller is sufficiently lower than the price of majority sellers. The consumers do not know the seller's race or the price charged before visiting them. In this environment, minority business owners decide the price of goods before the customer's visit. This decision is equivalent to the decision of whether they will sell to majority or not, since by lowering the price, minority business owners can sell to majority consumers. As the optimal decision, more productive minority business owners are more likely to lower the price than less productive minority business owners, since the opportunity cost of letting majority consumers pass by is too high. In other words, productive minority business owners attempt to capture a larger market, including majority consumers, by lowering the price of the product, since they produce more goods due to a higher production ability. As a result, the return to productivity among minority business owners is lower than the return among majority owners since more productive majority business owners do not have to lower their prices to capture a larger market.³

In the labor market of salary/wage (SW) sector, the minority workers are also discriminated by majority employers, but if their skills are discounted by the same fraction regardless of skill level, the return to skill is as same

³Consumers have inelastic demand below the reservation price in their model.

as one among majority workers. Holzer (1998) reported that larger firms unproportionately hire more African American workers than smaller firms even after controlling for firms' characteristics such as industry or location. He speculated stricter enforcement of anti-discrimination law for larger firms as an explanation for the finding. Since larger firms pay higher wages to workers, as far as larger firm select African American workers based on skill level, the return to skill in SW sector among African Americans can be even higher than among Caucasian Americans.

Because of the difference in the return to productivity as business owners, the selection into self-employment is different across racial (ethnic) groups; negative selection occurs more frequently among minority workers than among majority workers.

2.2 Credit Market Discrimination

Blanchflower, Levine, and Zimmerman (1998) analyzed discrimination in the small business credit market using the 1993 National Survey of Small Business Finances. They showed that African American business owners are twice as likely to have their business loan applications denied as Caucasian American business owners, even after controlling the factors that may affect credit-worthiness. They also reported that African American business owners are charged higher interest rates and are more likely to hesitate to apply for a business loan because they expect that they are likely to be rejected.

Since liquidity constraint is a critical issue in business start up and survival,⁴ discrimination in the small business credit market explains the lower rate of SE among African American workers.

This credit market discrimination also presumably creates negative self-selection into SE among African Americans when worker's ability and capital stock are complementary in the SE production. When ability and capital are complementary, it is optimal for the workers with high ability to run a big business. However, if a high ability African American who is a prospective SE worker learns that he cannot raise enough funds from the credit market, then he is likely to work as salary/wage worker because his high ability can be well-compensated without being constrained by the credit market. On the other hand, low ability African Americans who have a low level of optimal capital stock are more likely to finance their investments without being constrained by the credit market. The key here is that high ability African American workers are disproportionately penalized compared with low ability African Americans due to credit market discrimination when they attempt to be SE. As a result, negative self-selection into SE occurs among African Americans.

3 Previous Evidence

Both consumer discrimination and credit market discrimination predict negative self-selection into SE among African American workers. Using 1980

⁴See Holtz-Eakin, Joulfaian, and Rosen (1994a), Holtz-Eakin, Joulfaian, and Rosen (1994b), and Evans and Jovanovic (1989).

Census cross section data, Borjas and Bronars (1989) examined the differential self-selection rule into SE to test their own consumer discrimination hypothesis. Using Heckman's sample selection correction method, the earnings of each sector were broken down into (i) the part explained by background variables, (ii) the part due to self selection, and (iii) an error term correlated with neither background variables nor self-selection. The model they employed is laid out as follows. At first, the decision to be SE is estimated by the model:

$$se_i = 1(x_i\pi + v_i \geq 0), v_i \sim N(0, 1), \quad (1)$$

where x_i is a set of variables that determines a worker's SE status. Then the earnings in each sector is defined as

$$\ln w_i^{sw} = x_{1i}\beta^{sw} + e_i^{sw}, \quad (2)$$

and

$$\ln w_i^{se} = x_{1i}\beta^{se} + e_i^{se}, \quad (3)$$

where x_{1i} is a set of variables that affects a worker's earnings in each sector, which is a part of x_i . Since x_{1i} are included in the wage equations, the difference in the expected earnings in each sector affects workers' sector choice. In addition, the part of x_i not included in x_{1i} may include the factor that affects the decision to be SE but not earnings such as preference to autonomy or attitude toward risk. The error terms (v_i, e_i^{sw}, e_i^{se}) are independent of x_i with zero mean. The error terms in the wage equations are correlated with the error terms in the selection equation in the way that $e_i^{sw} = \gamma^{sw}v_i + u_i^{sw}$

and $e_i^{se} = \gamma^{se}v_i + u_i^{se}$. The conditional expectation of earnings in each sector after the selection is expressed as

$$E(\ln w_i^{sw}|x_i) = x_{1i}\beta^{sw} + \gamma^{sw}E(v_i|v_i < -x_i\pi), \quad (4)$$

and

$$E(\ln w_i^{se}|x_i) = x_{1i}\beta^{se} + \gamma^{se}E(v_i|v_i \geq -x_i\pi). \quad (5)$$

Here $\gamma^{sw} < 0$ implies positive self-selection into the salary wage sector (Those who become SW based on unobserved characteristics earn more as SW workers) and $\gamma^{se} > 0$ implies positive self-selection into the SE sector (Those who become SE based on unobserved characteristics earn more as SE workers). Borjas and Bronars (1989) implemented these two step estimations for each racial group using the Census 1980 file and found statistically significantly positive γ^{se} for Caucasians, but could not reject $\gamma^{se} = 0$ for African Americans. They interpreted these results as evidence of positive self-selection into SE among Caucasian Americans but not among African Americans.

However, without relying on the non-linearity of the inverse mill's ratio, the identification of β^j and γ^j , ($j = sw, se$) depends on the variable included in x_i but excluded from x_{1i} (Olsen (1980)). Borjas and Bronars (1989) used wife's education and other regional level variations⁵ as excluded variables. However this excluded variable assumption is rather tenuous, considering the possibility that the excluded variables may directly affect the earnings

⁵Specifically, the SMSA's unemployment rate, the population growth between 1970 and 1980, the crime rate, the level of local government expenditure, and the mean income and education levels in the local labor market were used.

of each sector. For example, suppose “ability” increases earnings but is not included as a control variable in the earnings equation and the “ability” is positively correlated with wife’s education conditioned on x_{1i} because of assortative mating.⁶ Since higher wife’s education encourages husband’s SE (the estimated coefficient in π corresponding to wife’s education was positive), “ability” and $E(v_i | v_i \geq -x_i\pi)$ are negatively correlated conditioned on x_{1i} because the smaller value of v_i satisfies the conditioning set. This negative correlation causes the downward bias for the estimate of γ^{se} . The degree of bias roughly depends on the strength of the correlation between “ability” and wife’s education, in the other words, the strength of assortative mating. If stronger assortative mating occurs among African Americans, then the downward bias of the estimate of γ^{se} is larger for African Americans. This example simply points out that the results obtained in Borjas and Bronars (1989) are attainable even in the absence of the different selection mechanisms across racial groups, but it in no way denies the results they obtained. Regional economic indicators, which were used as the other set of excluded variables in their study, are also likely to affect SE worker’s earnings, and the effect can differ across racial groups.

Considering the tenuousness of the identifying assumption employed in their study, their conclusion is suggestive rather than conclusive. Therefore, it is worth reexamining the selection mechanism across racial groups using a

⁶Assortative mating is the notion that individuals marry partners with similar socio-economic background characteristics. See Pencavel (1998) and Lewis and Oppenheimer (2000) for evidence of assortative mating by education.

different identification strategy.

4 The Identification of Self-Selection Using Panel Data

The difficulty in identifying the selection mechanism using cross section data arises because only data after selection is available. Panel data largely removes the restrictive assumption when identifying the self-selection mechanism because the data before the event of self-selection is available. In the current context, the earnings of salary/wage workers of two groups, the group of workers who become SE a year later and the group who stay in the salary/wage sector, offer rich information with which to identify the self-selection mechanism. Specifically, the distribution of $(\ln w_{it}^{sw}|Caucasian, se_{it+1} = 1)$ and $(\ln w_{it}^{sw}|Caucasian, se_{it+1} = 0)$ can be compared to see whether those who earn more as salary/wage workers self-select into SE in the following year. For example, if

$$E(\ln w_{it}^{sw}|Caucasian, se_{it+1} = 1) - E(\ln w_{it}^{sw}|Caucasian, se_{it+1} = 0) > 0 \quad (6)$$

is found, then positive self-selection into SE occurs at the mean of the salary/wage distribution among Caucasian workers. The difference in the salary/wage distribution can also be evaluated at several percentiles to see whether self-selection occurs at either the top or the bottom of the salary/wage distribution.

Comparing the self-selection mechanism across racial groups reveals dif-

ferences in this mechanism across groups. Specifically, examining

$$[E(\ln w_{it}^{sw}|Caucasian, se_{it+1} = 1) - E(\ln w_{it}^{sw}|Caucasian, se_{it+1} = 0)] \\ - [E(\ln w_{it}^{sw}|African, se_{it+1} = 1) - E(\ln w_{it}^{sw}|African, se_{it+1} = 0)] \quad (7)$$

reveals the differences in the selection mechanism at the mean of pre-selection wage distribution as salary/wage workers. The difference in distribution can also be examined at several percentiles to see if the racial differences are larger at higher percentiles of pre-selection wage distribution as predicted by consumer or credit market discrimination hypothesis. The idea is implemented by picking up several points of the distribution and estimating the following models. The selection at the mean is revealed through estimating the model:

$$E(\ln w_{it}^{sw}|aa_i, se_{it+1}) = \beta_0 + \beta_1 aa_i + \beta_2 self_{it+1} + \beta_3 aa_i \cdot self_{it+1}, \quad (8)$$

where aa_i takes one if individual i is African American. The rejection of $\beta_3 = 0$ implies different selection mechanisms at the mean between African Americans and Caucasian Americans. The selection at several percentiles can be examined through estimating the model:

$$P_j(\ln w_{it}^{sw}|aa_i, se_{it+1}) = \delta_{0j} + \delta_{1j} aa_i + \delta_{2j} self_{it+1} + \delta_{3j} aa_i \cdot self_{it+1}, \quad (9)$$

where P_j is j th percentile of the distribution. Again, the rejection of $\delta_{3j} = 0$ implies the difference in the selection mechanism at j th percentile of the distribution. Both consumer discrimination theory and credit market discrimination theory predict negative δ_3 for high j , since the theories predict,

that productive African Americans do not self-select into SE while productive Caucasian Americans do. Being able to examine the differential selection mechanism across the distribution in a straightforward way is the advantage of using panel data.

5 Data

The two-year panel created by matching adjacent years of the Current Population Survey (CPS) Out Going Rotation (ORG) group is used to implement the statistical analysis introduced in the previous section. Since African American SE workers are rare in the US work force, a panel containing a large number of observations is needed. Because of the relatively small number of cross-sectional units, conventional panel data such as the National Longitudinal Survey or the Panel Study of Income Dynamics are inadequate for this purpose. The merit of using matched CPS panel is its large sample size, which includes 50,000 households for each cross-section.

CPS randomly picks households according to its stratified sampling scheme and interviews selected households for four consecutive months. Then after a gap of 8 months, the surveyor returns to the original address and interviews for four consecutive months again. Because of this sampling scheme, the first four months of survey and the last four months of survey can be matched if the household does not move. In the fourth and eighth month (Out Going Rotation group), earnings-related information is collected in "earner study" section. In this study, individuals recorded in the CPS ORG files from 1990

to 2000 were matched with the data of following year using household id, household number, and within household id (line number). Then erroneous matches were checked using race, sex, and age information.⁷ Since household id were scrambled in June, July, and August 1995, the matches that included or crossed over these months were impossible. The sample construction is tabulated in Table 1. After conditioning based on the availability of valid job class (either private, government, or self-employed) information, age between 16-65, valid household id, linenumbr and household number, and the availability of non-imputed age, sex, and race information, there are 455107 observations available for the fourth month of the survey. Among these observations, 320649 observations were matched based on the household id, household number, line number, and race, sex, and age information. Thirty percent of the observations from the first year was lost; however, this rather high attrition rate is usual for matched CPS panel data, mainly due to moving households; CPS does not follow movers.⁸ Among those matched observations, male salary/wage workers with valid hourly rate of pay information in the first year of survey are selected. This sample restriction reduces the sample size to 235237 for 10 independent 2-year panels.

⁷See Madrian and Lefgren (2000) for the various methods of matching and their strengths and drawbacks.

⁸See Peracchi and Welch (1995) and Neumark and Kawaguchi (2001) for the usual attrition rate in matched CPS panels.

6 Results

Table 2 tabulates the share of SE workers in the total workforce for Caucasian Americans and African Americans. SE workers account for 16.03% and 6.23% of total work force among Caucasian Americans and African Americans, respectively. These numbers are slightly higher than the numbers calculated by Fairlie (2000b) (13 - 14 % and 5% for Caucasian and African, respectively, during the 1990s) or Borjas and Bronars (1989) (12% for Caucasian and 4.5% for African in 1980), but the SE rates for Caucasian Americans and African Americans are almost the same. The large difference in SE rate is partly due to the difference in the entrance rate in SE between Caucasian Americans and African Americans. The lower panel of Table 2 indicates that about 2.85% of salary/wage Caucasian American workers become SE in the following year, while the number for African Americans is about 1.59%. The data set used in this study confirms the lower rate of SE among African Americans and their lower transition rate into SE.

Table 3 tabulates the difference of observable characteristics of salary/wage workers between the two groups of workers. The first group consists of those who become SE and the second group consists of those who stay in salary/wage jobs in the second year. This table attempts to reveal self-selection based on observable characteristics. The notable difference in the selection across racial groups is that among Caucasian Americans, future SE workers are more likely to have higher educational background compared with

future stayers. While 33.5% of future SE workers have more than 16 years of education, only 27.4% of future salary/wage workers have that amount. This positive self-selection into self-employment based on higher educational attainment is not found among African American workers. The difference-in-differences estimate is marginally different from zero in the statistical sense, and the difference is large in its magnitude. On the other hand, there is not much difference in the age distribution among those who become self-employed across racial groups.

Figure 1 Panel A draws the distribution of hourly wages for two groups of Caucasian salary/wage workers: those who become SE in the following year and those who stay in salary/wage jobs in the following year. The mode of distribution of future SE workers is located left of future salary/wage workers, and this indicates negative self-selection into self-employment at the mode. The other notable difference is the fatter tail distribution of future SE workers compared with stayers. Both high earners and low earner as salary/wage workers are likely to become SE in the following year. This result suggests that examining central tendency is not sufficient when we discuss self-selection mechanisms into SE among Caucasian Americans.

Figure 1 Panel B repeats the same exercise using African American salary/wage workers as the analysis sample. Nearly the entire wage distribution of future SE workers is located to the left of the wage distribution of future salary/wage workers, and the distribution of future SE workers has a much fatter left tail. Positive self-selection into SE rarely occurs, and only at the high end of the

distribution.

Examination of Figure 1 roughly suggests the existence of both positive and negative self-selection into SE among Caucasian Americans, but almost only negative self-selection into SE among African Americans. To see this point more rigorously, the results of mean and quantile regressions are reported in Table 4. Slightly positive selection at the mean (0.018 with s.e. 0.009) is found among Caucasian Americans and negative self-selection (the difference from Caucasian is -0.124 with s.e. 0.038) into SE is found for African American workers. These results roughly match those obtained in Borjas and Bronars (1989), while they found positive self-selection for Caucasian American and no selection among African American. Results of the quantile regressions clearly indicate the two-sided self-selection among Caucasian Americans, but only negative self-selection among African Americans. Among Caucasian Americans, the 90 percentile of hourly wages for future SE workers is about 17.6% ($\exp(0.162) - 1$) higher than the same percentile of future stayers. Positive self-selection is found at 75 percentile and negative self-selection into SE is consistently found for lower percentiles. Among Caucasian Americans, positive self-selection into SE occurs at the upper tail of the distribution, while negative self-selection occurs at the lower tail of the distribution. However, positive self-selection balances with negative selection, and only slightly positive self-selection is found at the mean. On the other hand, only negative self-selection is observed among African Americans evaluated at any percentile. The difference in the selection mechanism

between Caucasian and African is stronger at the higher percentile of the distribution.

Since the analysis so far has not considered the difference in educational background and age to examine the self-selection mechanism, the observed difference in the self-selection mechanism among racial groups may be a product of differential observable background characteristics as we observed in Table 3. To draw a clearer picture of the self-selection mechanism based on unobservable characteristics, which is the main concern in the discussion of self selection, an analysis using the residual of the wage regression on education dummies, age, and its square appears in Figure 2. The distribution for Caucasian Americans that appears in Panel A clearly indicates two sided self selection. Future SE workers have a fatter tail distribution than future salary/wage workers. On the other hand, the distribution of hourly wage for future SE workers is always located to the right of the distribution for future salary/wage workers among African Americans. Table 5 contains the results of the mean and quantile regression with education dummies, age, and its square as independent variables. After controlling observed characteristics, negative self-selection into self-employment is found at the mean of the distribution for Caucasian Americans. Future SE workers earn about 3.7% less than future salary/wage job stayers, but even more negative self-selection is found among African Americans. Future self-employed African Americans earn about 6.3% less than future stayers. An examination of the quantile regression results reconfirms the two-sided self-selection among Caucasian

Americans and negative self-selection among African Americans. While the top 10% of future SE Caucasian American workers earns 10.0% more than the top 10% of future Caucasian American salary/wage workers, the top 10% of future SE African Americans earns 12.2% less than the top 10% of future salary/wage African American workers. At the same time, while the bottom 10% of future SE Caucasian American workers earns 14% less than the bottom 10% of future Caucasian American salary/wage workers, the bottom 10% of future SE African Americans earns only 3.3% less than the bottom 10% of future salary/wage African American workers. The results based on education-age adjusted wage clearly indicate both positive and negative self-selection into SE among Caucasian Americans but uniformly negative self-selection among African Americans.

As seen so far, the analysis based on panel data draws a richer picture of the self-selection mechanism than the analysis using cross-sectional data. The evidence from the panel data at least does not contradict with the hypothesis of lower return to skill among African American SW workers. The interpretation of the negative self-selection into self-employment among African Americans cannot be attributed to the specific hypothesis that predicts lower return to skill among African American SE workers, but the evidence is hard to interpret without presuming some sort of discrimination against eligible African American SE workers that is harsher than the discrimination against eligible African Americans SW workers. The return to skill among African Americans can be higher in the SW sector than among

Caucasian Americans if skilled SW African American workers are less discriminated than unskilled SW African American workers due to possible non-uniform enforcement of anti-discrimination policy across skill groups in SW sector as implied by Holzer (1998).

The evidences found in this paper indirectly support the hypothesis that predicts relatively lower return to skill in SE sector than in SW sector among African American. Accordingly, both consumer or credit market discrimination combined with anti-discrimination law enforcement in salary/wage sector survive as explanations for the lower SE rate among African Americans.

7 Conclusion

This paper attempted to shed light on the mechanism behind the lower rate of SE among African Americans compared with Caucasian Americans. Several theories of discrimination have been explored to explain the lower rate of SE among African Americans. Among the theories, both consumer discrimination, discrimination in the credit market combined with anti-discrimination law enforcement in SW sector explain why African American workers, eligible African American workers in particular, are less likely to be SE.

The previous influential study by Borjas and Bronars (1989) found positive self-selection into SE among Caucasian Americans and neutral self-selection into SE among African Americans using Heckman's sample selection correction method. A similar but slightly different conclusion was obtained in this paper under a rather weak identifying assumption using matched CPS

panel data. Current salary/wage workers were divided into prospective SE workers and prospective salary/wage workers and the current wage of these two groups were compared. An examination of entire distribution revealed that both high earners and low earners self-select into SE among Caucasian Americans. However, among African Americans, only current low earners select into SE. The data draw sharp pictures of two-sided selection into SE among Caucasian American workers and negative selection into SE among African Americans.

The contribution of this paper is twofold. First, an alternative identification strategy using panel data confirmed the results obtained in a previous study. Second, the examination of whole distribution revealed the complex two-sided self-selection mechanism among Caucasian Americans. This complex mechanism could not have been found if only the mean of the distribution had been examined, as in Heckman's method.

The finding obtained in this paper is consistent with several hypothesis predicting lower rate of return to skill in SE sector among African workers. Although the specific hypothesis that predicts it cannot be pinned down, the evidences found in this paper certainly offer a step stone for the further investigation into the reason why there are so small number of African American owned small businesses. Pinning down the specific hypothesis using a clever identification strategy with richer panel data would be a challenging but rewarding future project.

References

- Blanchflower, D. G. P. B. Levine and D. J. Zimmerman, 1998, Discrimination in the Small Business Credit Market, NBER Working Paper 6840.
- Borjas, G. J. and S. G. Bronars, 1989, Consumer discrimination and self-employment, *Journal of Political Economy*, 97(3), 581-605.
- Evans, D. and B. Jovanovic, 1989, An Estimated Model of Entrepreneurial Choice under Liquidity Constraints, *Journal of Political Economy*, 97(4), 808-27.
- Fairlie, R. W., 1999, The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment, *Journal of Labor Economics*, 17(1), 80-108.
- , 2000a, Earnings Growth among Young Less-Educated Business Owners, Mimeo, University of California, Santa Cruz.
- , 2000b, Recent Trends in Ethnic and Racial Self-Employment, Mimeo, University of California, Santa Cruz.
- Fukuyama, F., 1995, *Trust: The Social Virtues and the Creation of Prosperity*. Free Press, New York.
- Holtz-Eakin, D. D. Joulfaian and H. Rosen, 1994a, Entrepreneurial Decisions and Liquidity Constraints, *Rand Journal of Economics*, 25(2), 334-347.

- , 1994b, Sticking It Out: Entrepreneurial Survival and Liquidity Constraints, *Journal of Political Economy*, 102(1), 53–75.
- Holzer, H. J., 1998, Why Do Small Establishments Hire Fewer Blacks than Large Ones, *Journal of Human Resources*, 32(4), 897–914.
- Kim, G. and G. Deltas, 2002, The Dynamics of Self-Employment over the Life-Cycle: An Entry-Exit Decomposition Analysis, University of Illinois.
- Lewis, S. K. and V. K. Oppenheimer, 2000, Educational Assortative Mating across Marriage Markets: Non-Hispanic Whites in the United States, *Demography*, 37(1), 29–40.
- Madrian, B. C. and L. J. Lefgren, 2000, A Note on Longitudinally Matching Current Population Survey (CPS) Respondents, *Journal of Economic and Social Measurement*, 26, 31–62.
- Moore, R. L., 1983, Employer Discrimination: Evidence from Self-Employed Workers, *The Review of Economics and Statistics*, 65(3), 496–501.
- Neumark, D. and D. Kawaguchi, 2001, Attrition Bias in Economic Relationships Estimated with Matched CPS Files, NBER Working Paper 8663.
- Olsen, R. J., 1980, A Least Squares Correction for Selectivity Bias, *Econometrica*, 48(7), 1815–1820.
- Pencavel, J., 1998, Assortative Mating by Schooling and the Work Behavior

of Wives and Husbands, *American Economic Review, Papers and Proceedings*, 88(2), 326–329.

Peracchi, F. and F. Welch, 1995, How Representative are Matched Cross-Sections? Evidence from the Current Population Survey, *Journal of Econometrics*, 68(1), 153–79.

Table 1: Sample Construction

Observations in 1990-2000, excluding the period of identifier renumbering	2227016
Male, Caucasian, or African American, age 16-65	521635
Work in private, government, or self-employed sectors	462079
All information used in consistency check is available and not assigned	455107
Naïve match	320649
Sex, race, and age are consistent across surveys	316326
Not self-employed & eligible for earner study in the first year	245617
Hourly rate of pay available in the first year	235237

Sample: Matched CPS files Outgoing Rotation Group

(1990-1991, 1991-1992, 1992-1993, 1993-1994, 1994-1995 (January-May), 1995-1996 (September-December), 1996-1997, 1997-1998, 1998-1999, 1999-2000)

Note: Household IDs were revised in June to August 1995. Observations in these months cannot be matched with the observations in the other months.

Adjacent files are matched based on hhid, linenummer, hhno, race, sex, and age. Age is allowed to be different between -1 and 3.

African and Caucasian men are included in the sample.

Table 2: Self Employment Rate and Transition into Self Employment by Racial Groups

	Caucasian	African	Caucasian -African
Number of Self Employment	46933	1472	
SE Rate	16.03	6.23	9.80 (0.17)
Total N	292704	23622	316326
Number of Transition from SW to SE (Annual)	6456	310	
Transition rate from SW to SE (Annual)	2.85	1.59	1.26 (0.10)
Total N	226152	19465	245617

Note: Standard errors are in parenthesis for the difference of the ratio between racial groups. The sample used for this table is different from the analysis sample for the regression analysis. The sample size is larger because the sample is not conditioned on the status of salary/wage worker and the availability of hourly rate of pay in the first year.

Table3: Descriptive Statistics Based on Year 1 Information

	Caucasian			African			(SE-SW) _C - (SE-SW) _A
	SW in Year 2	SE in Year 2	SE-SW	SW in Year 2	SE in Year 2	SE-SW	
Educ							
<12	0.127	0.115	-0.012 (0.004)	0.178	0.167	-0.011 (0.022)	-0.001 (0.023)
12	0.346	0.304	-0.043 (0.006)	0.400	0.394	-0.006 (0.030)	-0.037 (0.031)
13-15	0.252	0.246	-0.006 (0.006)	0.258	0.260	0.002 (0.027)	-0.009 (0.028)
16+	0.274	0.335	0.061 (0.006)	0.164	0.178	0.015 (0.024)	0.046 (0.024)
Age							
16-25	0.138	0.090	-0.048 (0.004)	0.131	0.104	-0.027 (0.019)	-0.021 (0.038)
26-35	0.261	0.253	-0.007 (0.006)	0.260	0.290	0.030 (0.028)	-0.037 (0.028)
36-45	0.295	0.326	0.031 (0.006)	0.296	0.279	-0.017 (0.028)	0.048 (0.028)
46-55	0.212	0.216	0.004 (0.005)	0.221	0.216	-0.006 (0.025)	0.009 (0.026)
56-65	0.094	0.115	0.021 (0.004)	0.091	0.112	0.020 (0.019)	0.001 (0.020)
HRP	13.159 (8.208)	14.241 (10.749)	1.082 (0.142)	10.454 (7.724)	9.727 (6.952)	-0.726 (0.428)	1.808 (0.451)
N	210863	5796	216659	18309	269	18578	235237

Table 4: Selection Based on Hourly Rate of Pay in Year 1

	Mean of Log Hourly Wage	10 %tile of Log Hourly Wage	25 %tile of Log Hourly Wage	Median of Log Hourly Wage	75 %tile of Log Hourly Wage	90 %tile of Log Hourly Wage
SE in Year 2	0.018 (0.009)	-0.012 (0.000)	-0.025 (0.000)	-0.000 (0.000)	0.077 (0.004)	0.162 (0.005)
African	-0.222 (0.004)	-0.118 (0.000)	-0.223 (0.000)	-0.248 (0.000)	-0.233 (0.002)	-0.260 (0.003)
SE in Year 2 × African	-0.124 (0.038)	-0.052 (0.000)	-0.040 (0.000)	-0.132 (0.000)	-0.182 (0.019)	-0.265 (0.023)
Constant	2.422 (0.006)	1.638 (0.000)	2.029 (0.000)	2.427 (0.000)	2.803 (0.003)	3.121 (0.004)
No. of observations	235237	235237	235237	235237	235237	235237
R ²	0.01	-	-	-	-	-

Note: Hourly rate of pay is deflated by the Consumer Price Index (all urban consumers, U.S. city average, all items). Year dummies are included.

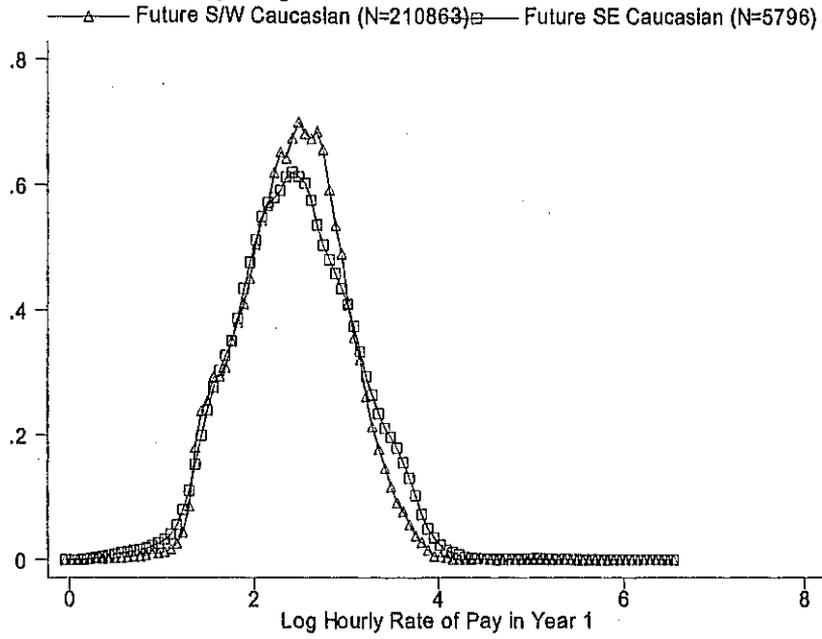
Table 5: Selection Based on Hourly Rate of Pay in Year 1, Control Variables Adjusted

	Mean of Log Hourly Wage	10 %tile of Log Hourly Wage	25 %tile of Log Hourly Wage	Median of Log Hourly Wage	75 %tile of Log Hourly Wage	90 %tile of Log Hourly Wage
SE in Year 2	-0.037 (0.008)	-0.142 (0.011)	-0.089 (0.008)	-0.032 (0.006)	0.034 (0.007)	0.100 (0.009)
African	-0.172 (0.004)	-0.172 (0.006)	-0.192 (0.005)	-0.179 (0.004)	-0.159 (0.004)	-0.151 (0.005)
SE in Year 2 × African	-0.090 (0.035)	-0.033 (0.051)	-0.031 (0.039)	-0.119 (0.029)	-0.117 (0.034)	-0.122 (0.041)
Education < 12	-0.232 (0.003)	-0.209 (0.006)	-0.197 (0.004)	-0.218 (0.003)	-0.238 (0.004)	-0.228 (0.005)
Education 13 – 15	0.107 (0.003)	0.060 (0.004)	0.081 (0.003)	0.109 (0.003)	0.121 (0.003)	0.131 (0.004)
Education > 16	0.412 (0.003)	0.319 (0.004)	0.393 (0.003)	0.434 (0.003)	0.458 (0.003)	0.471 (0.004)
Age	0.079 (0.001)	0.067 (0.001)	0.074 (0.001)	0.082 (0.001)	0.088 (0.001)	0.086 (0.001)
Age2	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)
Constant	0.555 (0.011)	0.337 (0.021)	0.399 (0.015)	0.485 (0.012)	0.595 (0.013)	0.852 (0.016)
No. of observations	235237	235237	235237	235237	235237	235237
R ²	0.33	-	-	-	-	-

Note: Hourly rate of pay is deflated by the Consumer Price Index (all urban consumers, U.S. city average). Year dummies are included. High school graduate is the base category of the educational dummies.

Figure 1: Distribution of Hourly Rate of Pay as Salary Wage Workers in Year 1

Panel A: Caucasian Salary/Wage Workers



Panel B: African Salary/Wage Workers

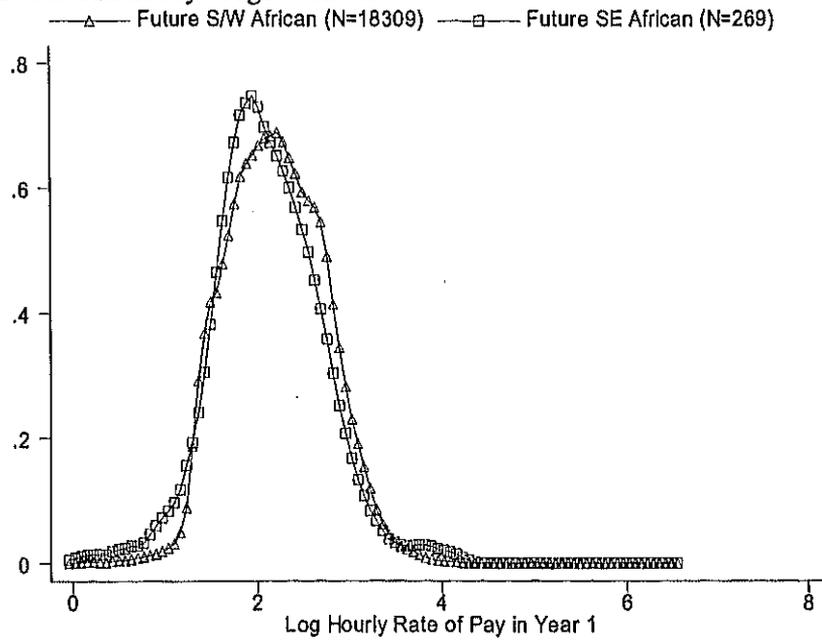
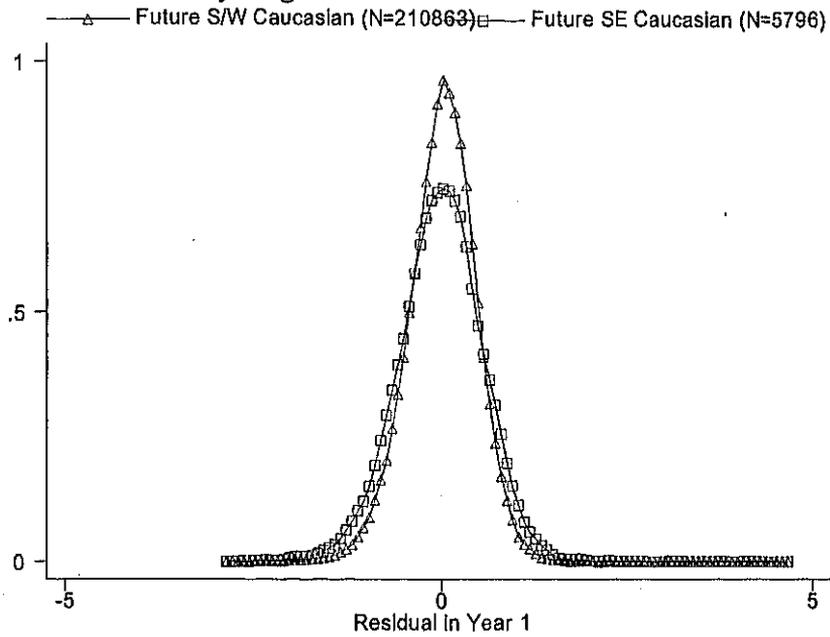
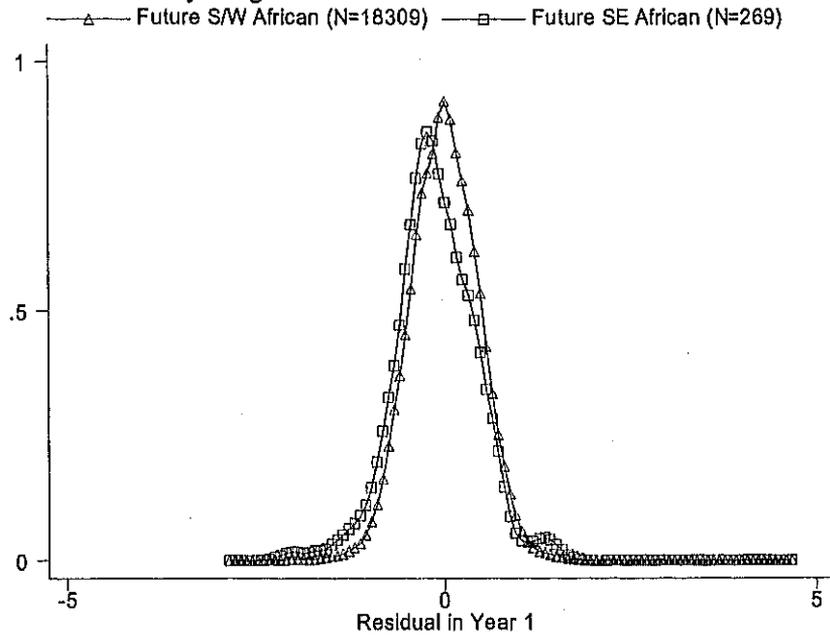


Figure 2: Distribution of OLS Residual of Wage Regression in Year 1

Panel A: Caucasian Salary/Wage Workers



Panel B: African Salary/Wage Workers



Note: The residual is the residual of the following OLS regression.

$$\ln w = \beta_0 + \beta_1 educ<12 + \beta_2 educ12 + \beta_3 educ13-15 + \beta_4 educ>16 + \beta_5 age + \beta_6 age2 + u.$$

The sample consists of salary / wage workers and the models are estimated separately for Caucasian and African workers.