1-1-2011

# Three Essays on Labor and Personality

Nidhi Pande
*Washington University in St. Louis*

WASHINGTON UNIVERSITY

Department of Economics

Dissertation Examination Committee:
Barton Hamilton, Chair
Sebastian Galiani
Juan Pantano
Bruce Peterson
Bob Pollak
Carl Sanders

# Three Essays on Labor and Personality

by

Nidhi Pande

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2011

Saint Louis, Missouri

# Dissertation Abstract


Three Essays on Labor and Personality

by
Nidhi Pande

Doctor of Philosophy in Economics

Washington University in St. Louis, 2011

Professor Barton Hamilton, Chair



The scope of the dissertation is microeconometrics. The first essay is on *human capital formation*, the second essay is on *personality*, and the third essay is on *labor market decisions*.

Using a randomized experiment, the first essay examines the impact of mother's human capital on the cognitive and non cognitive skills of her preschool children.

The second essay[1] examines the impact of the big five personality traits on the decision to be self employed and on the income of salaried vs. self employed people. We try to distinguish the impact of personality traits on labor market performance from the relationship between personality and preferences for entrepreneurship.

Finally, in the third essay[2] we are trying to estimate the labor market wage premium for shift workers. We use an equilibrium sorting framework to model location decisions around the clock. Using the estimated model we try to disentangle the amenity value of daylight from social interaction effects.


[1]This essay is a joint work with Barton Hamilton.
[2]This essay is a joint work with Juan Pantano.

# Acknowledgements

I want to acknowledge many people for the successful completion of my thesis.

I begin by thanking Bart who is instrumental in my completing my thesis. Discussing ideas with Bart was always a pleasure and he always listened to all my ideas/apprehensions patiently. I nearly left the program during my second year till I found Bart who started me off on a project and I stayed on. I want to thank Bob for the Work and Public Policy seminar series and the numerous discussions I have had with him. A big thanks to Juan for putting up with tons of mistakes I made during the work on our paper and teaching me a lot about how things work. Also thanks to Sebastian Galiani, Carl Sanders and Bruce Peterson for being on my committee. Finally among the faculty I would like to thank John Nachbar for helping me pull through it during difficult times.

Thanks to Nick and Carmen, my classmates for always being there for me. We did it guys. Lots of gratitude to Piyusha my friend for making me think through organising the whole defense and making me believe that I could finish it. Thanks to Anubha, Divya, Adhira, Karthik.

Finally a big thanks and hug to my father, mother, sister, brother, Jojo and Rajeswari . I would not have been able to finish without your love support and best wishes.

# Contents

# List of Tables

# List of Figures

Chapter 1

# Assessing the Impact of Mothers' Investment in Human Capital on Childrens' Outcomes

## 1.1 Introduction

The view that cognitive,social and behavioral skills are important determinants of an individual's academic and socioeconomic success in life is gaining ground in the literature studying academic and economic inequality (Cawley Heckman and Vytlacil 2001; Herrnstein and Murray 1994; Murnane Willett and Levy 1995; Neal and Johnson 1996; Bowles Gintis and Osborne 2001; Heckman and Rubinstein 2001; Heckman Stixrud and Urzua 2006). Cognitive skills refer to mental or intellectual abilities of an individual while social and behavioral skills refer to what is known as emotional intelligence. A widely held view in this literature is that these abilities are largely influenced by pre-school /early childhood experiences of the child. A large multidisciplinary literature has been trying to determine the impact of parental characteristics, early home environment and school quality in producing these skills but there is still a lack of consensus on the relative importance of these factors on skills. Determining the relative importance of these factors coupled with the right age at which they can be influenced becomes particularly important for examining the efficacy of various policies targeting child development. Our paper is a step in this direction.

Using a randomized experiment that provided education, training and life skill services to teenage high school dropout mothers', we try to ascertain the impact of an increase in the mothers' human capital on the cognitive and behavioral skills of their preschool children. There have been very few attempts to examine the impact of an increase in the mothers' human capital on the abilities of children. To the best of our knowledge this is one of the the first papers that is trying to examine the impact of an increase in mothers' human capital on preschool children's cognitive and emotional skills.[1] We also examine the importance of other factors considered in literature like the home environment and maternal employment. Maternal employment can increase the child's abilities because it may increase the human capital the mother passes on to the child. It might also work through the income effect route. It may also reduce the development because employment might conflict with the time the mother can give to the child. These questions have been examined earlier not only in economics but in child development, psychology and health literature also (Desai et al 1989; Brooks-Gunn et al 2002; Auld and Sidhu 2005; Murnane et al 1981; Carlson and Corcoran 2001).

Most of the previous studies in economics have tried to estimate child skill development through education/skill production functions. Endogeneity of explanatory variables due to missing data on children's home and school inputs as well as missing genetic endowments coupled with measurement errors has plagued estimation of these production function parameters. Additionally, different specifications of

[1]We recently became aware of work by Magnuson (2007) on similar lines on a different dataset

the function results in varied conclusions about the outcomes (Krueger 2003). The main problem in estimating the effect of mothers' human capital on the child's skills is that mothers' unobserved abilities will influence both her and the child's skill acquisition thereby giving an upward bias to the estimated impact of mothers' skills. Our paper gets around this issue by virtue of the randomization in the data. The increase in the home inputs, education and skills of the mothers' in the experimental group is due to their participation in the new chance program and hence being in the treatment group serves as an instrument for the endogenous variables.

Nearly all existing work examining child cognitive development include mothers' education as an input while estimating the cognitive/ non cognitive achievement production function. In an early estimation of the production function for human capital of children Leibowitz (1974) concludes that even in a sample of high IQ children mothers' education was significantly related to child's IQ.Cunha and Heckman (2008) estimate models of evolution of cognitive and non cognitive skills over the life cycle of children. They find that parental inputs affect the formation of both non cognitive skills and cognitive skills. Direct measures of mothers' ability affect cognitive skills but not non cognitive skills. They also find that cognitive skills are shaped at earlier ages while non cognitive skills are more malleable at later ages. They claim that ages 6-7 are the sensitive periods for cognitive skill formation while 8-9 years are the sensitive periods for non cognitive skills. Our paper examines the sensitivity of these skills at earlier ages thereby eliminating the need to separate out effects of school enrollment and increase in maternal human capital.

An important difference between experimental and non experimental studies is pointed out by Todd and Wolpin (2003). Production function estimation falls under the non experimental category because here inputs in the production function are subject to choices made by parents and school. Experimental studies on the other hand involve at least some or all the inputs being chosen by random assignment. They point out that the parameters estimated in experimental/natural experiment studies typically differ from those estimated in non experimental studies and one type of evidence does not substitute for the other. They warn against drawing comparisons between the two types of estimates under the presumption that they are estimating the same parameter. Our work is closely related to Rosenzweig and Wolpin (1994). They examine the effect of maternal education on the intellectual achievement of the child. They set out a model incorporating human capital production in children, fertility and maternal schooling investment as well as heterogeneity in human capital endowments. Their results indicate that maternal schooling attainment has positive influence on the achievement measures of children but not on their ability measures, net of other inputs and endowment heterogeneity among mothers'. They also find that mothers' who remain in school after having a child do not augment the intellectual growth of that child but do augment the intellectual growth of subsequent children. Our work is different from theirs in two aspects: (i) they examined the returns to maternal schooling for school going children while we do it for preschool children and (ii) they used a non experimental dataset (NLSY) while we do the estimation for an experimental dataset. Our results are similar to what they find. An increase in mothers' human capital does not seem to significantly affect the cognitive ability of the child born before the increase.

The rest of the paper is organized as follows. Section 2 covers the data. Section 3 talks about the methodology and results. Section 4 concludes.

## 1.2 Data

The data for our study comes from the New Chance project. New Chance was a voluntary demonstration project that provided comprehensive education, training and other services intended to increase the long-term self-sufficiency and well being of a group of high school dropout teenage mothers' who were receiving Aid to Families with Dependent Children (AFDC). During the program's demonstration phase, which began in 1989 and concluded in 1992, New Chance was operated by community-based organizations, schools, a community college, and municipal agencies at 16 locations (or "sites") in 10 states across the country.[2]

It was targeted at 16 to 22 year old mothers' who had first given birth at 19 or younger, were not pregnant when they entered the program, had dropped out of high school and were receiving cash welfare assistance. Most women enrolled in the program voluntarily, though some were referred by welfare-to-work programs. Women who applied and were determined to be eligible for New Chance were randomly assigned to one of two groups: the experimental group or the control group. The experimental group could enroll in the program while the control group could not join New Chance but could receive other services available in their communities. New Chance was implemeted in two phases:

- Phase 1 centered on education, career exposure, and a number of services

[2]The New Chance program was run at the following 16 sites: Allentown (Pennsylvania), Bronx (New York), Chicago Heights (Illinois), Chula Vista (California), Denver (Colorado), Detroit (Michigan), Harlem (New York), Inglewood (California), Jacksonville (Florida), Lexington (Kentucky), Minneapolis (Minnesota), Philadelphia (Pennsylvania), Pittsburgh (Pennsylvania), Portland (Oregon), Salem (Oregon), San Jose (California). The distribution across the sites is presented in Table 1.8

falling under the general rubric of "personal development" (for example, parenting, family planning, and life skills). During this phase, services were delivered mostly at the program site. Typically, the program ran from 9 a.m. until 3 p.m. five days a week, with daily attendance at all classes expected. Local programs were intended to be small in size, enrolling 100 participants over 12 to 18 months and serving about 40 participants at any given time, in order to promote an intimate and personal environment in which participants and staff could establish close bonds.

- Phase 2 services encompassed occupational skills training and work experience (both of which were generally off-site) and ultimately job placement assistance. Although college was not a formal part of the New Chance model, staff members at some sites encouraged participants to enroll in college, especially in two-year programs with a vocational focus.

Enrollees were permitted to remain in the program for 18 months, throughout which time case managers were expected to counsel them and monitor their progress. Each site had case managers who kept track of each participants progress and provided continuous guidance and support. There were follow-ups at 18 and 42 months.

The outcome variables considered for measuring cognitive and non cognitive skills are Bracken Basic Concept Scale School Readiness Component (BBCS) and Behavior Problems Index (BPI) respectively. The BBCS is a measure of receptive language that assesses the mastery of basic concepts; the School Readiness Component consists of five subtests of the BBCS: colors, letter identification, numbers, comparisons, and shapes. The scores shown are standard scores on a scale that ranges from 1 to 19; a standard score of 6.9 corresponds to about the 15th percentile nationally. The BPI is a widely employed scale for describing the incidence of behavioral problems of children aged four or older, usually as described by a parent. Raw scores for the BPI and its six subtests were converted to standardized normed scores, which are based on data from the 1981 National Health Interview Survey. These standard scores (with a mean of 100) are standardized separately for boys and girls within single years of age. A higher score points to more behavioral problems. The potentially endogenous variables that we instrument for by using participation in the experiment include HOME scale and mothers' educational status and mothers' earning at month 42. The HOME (Home Observation Measurement of the Environment) scale is a survey measure of parenting and the home environment. It appraises the orderliness, cleanliness, and safety of the physical environment, the regularity and structure of the family's daily routine, the amount of intellectual stimulation available to the child and the degree of emotional support provided by parents. It does this through a combination of questions asked of the parent and items to be completed by the interviewer after spending time in the home observing the child's physical surroundings and the parent and child interacting with one another.

Our sample consists of 2079 women. Out of 88 were missing the BBCS, 35 were missing the BPI score and 237 were missing both. Also a few were missing some of the explanatory variables. Hence our final estimation includes 1754 observations for BBCS and 1807 observations for BPI. The BBCS and BPI scores for the sample are summarized in Table 1.1. 61 were missing home score but we did not drop these observations, we included a missing home score dummy in the estimation. Basic summary statistics of the sample used are presented in Table 1.1. The women were on an average 19 years old when they joined the program. They had given birth when they were around 17 years of age. More than 50% of the enrolees were black and 60% of the program participants had completed grade 10 or less. We also did some mean comparison tests to check the randomness of the sample at baseline. Results are presented in Table 1.2. Since mothers' were assigned to one or the other group at random, the two groups did not differ at the onset of the study. Therefore, any differences between them that emerged during the follow-up period can be attributed to the program. The distribution of sample across sites is provided in appendix B.

## 1.3 Methodology and Results

We use the following specification of cognitive achievement/behavior, where we include both current and past inputs:

$$Y_{ij}^s = \alpha_o + \alpha_1 T_{ij}^s + \alpha_2 F_{ij}^s + \delta_j + \epsilon_{ij} \tag{1.3.1}$$

where $s = c, nc$ is the cognitive and behavioral test score for the child, $Y_{ij}^s$ is outcome s for child i at location j, $T_{ij}^s$ is the vector of potentially endogenous inputs. Includes mothers' education, earnings and HOME score, $F_{ij}^s$ is the demographic chracteristics and household inputs at baseline. Includes mothers' age, education of grandparents, mothers' depression score. $\delta_j$ are the site fixed effects.

We begin by estimating the outcome equations by simple OLS. The results are presented in columns 1 and 3 of Table 1.3. Site dummies are included in both regressions to account for the differences across sites in the composition of women who enrolled for the program. We also use mothers' earning as a channel through which an increase in mothers' human capital can increase children's outcomes in our regression. Mothers' earning could impact child skill formation through two channels: (a) Mothers' employment results in income which in turn determines certain inputs that go into the child skill production function. Additionally the human capital/social skills acquired by the mother at work could also be beneficial to the child and hence there could be a positive impact of mothers' employment on the child skill formation (b) Employment could result in less time being spent with/on the child by the mother. It could also lead to depression leading to lesser inclination to provide quality parenting to the child. Other covariates include mother more than grade 10 at baseline, mothers' test of adult basic education(TABE) score at baseline, age of mother at baseline, indicator for black or hispanic, mothers' dad stayed with her at age 14, mothers' family never on welfare when young, mother ever married at baseline, childs' age is greater than 18 months at baseline, indicator for boy, at least one grandparent has completed high school or more, mothers' CESD depression score at baseline along with site dummies.

The sign on the coefficients from least squares estimation of the treatment variables are what we expect for both regressions. Mothers' education and home score has a positive impact on both cognitive and behavior measures, as does the earning in last 12 months by mother variable. For the cognitive skill score all of the treatment variables except mothers' earning in the last 12 months are significant at 1%. BPI on the other hand seems to be affected only by the home score.[3] The children of mothers' who have completed high school/GED by month 42 on an average have a BBCS score .47 higher than the score of ones whose mothers' do not have a high

[3]For BPI a higher score signifies a worse outcome

school/GED. If mother has some college credit by month 42 then the score gap increases by .69 (10% of the average). For a one unit increase in HOME score we see the BBCS score go up by .05 and the BPI score go down by .24. Even though the results conform to our beliefs about the relationship between these inputs and child skills, we approach them with caution because of the potential bias in the coefficients of the treatment variables due to omitted mothers' ability which we expect to be correlated with both the treatment and the dependent variable.

To address this problem we do IV estimation of our equations. Due to the randomized nature of the experiment we will expect mothers' in the treatment group to have higher levels of human capital due to their participation in the program and not due to their ability. Hence it can serve as an instrument for the endogenous variables. Interactions between site dummies and program participation also cause additional variation in the mothers' human capital due to the possible differences in the way the program was run across different sites by the case managers and program staff.[4] Hence, we instrument for the treatment variables with participation in the program and program participation interacted with the site dummies.[5] Table 1.3 columns 2 and 4 present the IV estimates for the cognitive and emotional skill equations. These coefficients tell us the impact of the increment in mothers' human capital on the child's skill formation instead of the total impact of mothers' abilities on the child's abilities. We control to some extent for the mothers' existing stock of skill by including the mothers' TABE score at baseline.

After instrumenting for the endogenous variables the impact of nearly all the

[4]Using multisite programs to create instruments is fairly common in literature. See Bloom (2005)

[5]Details provided in appendix B

treatment variables becomes insignificant. None of the treatment variables has a significant impact on the BBCS score. Additionally the signs associated with some of the coefficients also change even though they are now insignificant. Note that the mothers' has high school/GED dummies now have 'wrong' signs in both BBCS and BPI regression. The coefficent on mothers' earnings also changes signs for the BBCS score but is still insignificant. For both the BBCS and BPI score the significance of the HOME score goes away. Table 1.3 columns 2 and 4 seem to suggest that after controlling for family background, post birth investment in mothers' human capital is not significantly affecting the pre school outcomes of their child on average. However these results are averages over all children and could be hiding differences across groups. In order to examine this we estimate seperate regressions for race and age of child at baseline. This is done to check for differences in outcome for one particular race that might be driving the result. Since the data does not have the actual age of the child, some of the children might already be in school when the tests were administered. This might distort the results, hence we separately estimate the equations for children $< 18$ months at baseline (hence they will still be preschool at the time of taking the tests) and children $> 18$ months at baseline (might already be in school at the time of test). Results for race are presented in Table 1.4 panels A, B and C. There are some differences across racial groups in terms of the importance and significance of the treatment variables. The home score is consistently significant and positively related to the BPI score in the least squares regression across all racial groups except for the BBCS score in hispanics. All variables have the 'right' or expected sign in the OLS regression for blacks and hispanics. The mothers' earning coefficient has a 'wrong' sign for whites. The instrumental variables estimates vary in sign, magnitude and significance across the three groups. Home score remains significant at 5% and its impact increases the BBCS score in blacks. For whites and hispanics none of the variables are significantly determining the child outcomes once we account for endogeneity with mothers' ability. Table 1.5 panels A and B present the OLS and IV results for children $>$ 18 months and children $<$ 18 months at baseline respectively. The least squares outcomes for children $<$ 18 months are very similar to the overall results in terms of signs and significance of the coefficients. For children $>$ 18 months, home score is the only significant variable for both BBCS and BPI. After instrumenting mothers' education variables loose their significance and become of the 'wrong' sign for either BPI or BBCS across both categories in Table 1.5 panel B columns 2 and 4.

Hence instrumenting the potentially endogenous variables is giving insignificant and counterintutive results. This however does not imply that mothers' human capital does not impact the childs' outcomes. The significant increase in standard errors in panel B for all regressions raise suspicion about weak instruments, hence we checked the joint significance of the instruments in the first stage regressions. The F statistics associated with the first stage of the instrumental variable regressions are presented in Table 1.9. They confirm our suspicion of weak instruments. We then try to see the effectiveness of the program for mothers' and the impact on the children of the program through other measures.

### 1.3.1 Intent to Treat and Treatment on the Treated

We now look at the whether the program had any effect on the childs' outcome through a simple intent to treat (ITT) regression of the child outcome on the program assignment indicator.

$$Y_{ij} = \pi_o + \pi_1 Z_{ij} + \pi_2 X_{ij} + \delta_j + \epsilon_{ij} \tag{1.3.2}$$

where $Y_{ij}$ is outcome of interest for person i at location j, $Z_{ij}$ is the program assignment indicator, where

$$Z_{ij} = \begin{cases} 1 & \text{if mother in the treatment group} \\ 0 & \text{if mother in the control group} \end{cases}$$

$X_{ij}$ is the vector of exogenous inputs. Includes mothers' age, education of grandparents, mothers' depression score. $\delta_j$ are the site fixed effects.

We also do the above for the six subcomponents of the BPI score (anxious or depressed, antisocial, dependent, headstrong, hyperactive and peerconflict) to see if any one dimension was particularly affected. The analysis is also done for the Positive Behavior Index (PBI) and its subcomponents (autonomy, compliance/self control and social competence/sensitivity). Since by the second interview some children have started going to school, we also include a school performance variable in our ITT. In this variable the mother ranks the performance of the child in school on a scale of 0 to 10 with higher magnitude implying better performance. Additionally we also look at the effect of the program on the endogenous variables considered in the IV regressions above. The results are presented in column 1 of Table 1.6.

Columns 2 and 3 of Table 1.6 present estimates of the treatment on the treated (TOT) effects. For the TOT we generate the program participation indicator $D_{ij}$, where

$$D_{ij} = \begin{cases} 1 & \text{if participation hours in New Chance} > 0 \\ 0 & \text{if participation hours in New Chance} = 0 \end{cases}$$

Column 2 estimates the impact of the program by a simple regression of the outcome on the program participation indicator. TOT instruments the program participation indicator with the program assignment. Hence TOT estimates the equation:

$$Y_{ij} = \gamma_o + \gamma_1 D_{ij} + \gamma_2 X_{ij} + \delta_j + \epsilon_{ij} \tag{1.3.3}$$

where

$$D_{ij} = \beta_o + \beta_1 Z_{ij} + \beta_2 X_{ij} + \delta_j + \eta_{ij} \tag{1.3.4}$$

From all the three analysis it seems that the New Chance program is affecting the children outcome negatively. The coefficients of the program assignment and participation dummy indicate that it decreases the BBCS score and increases the BPI score for children. The coefficients for BPI are significant also. ITT coefficient indicates that the children of the treatment mothers' have a 3% lower BBCS score than the control group mothers'. The TOT also has similar coefficient even though both are insignificant. On examining the subcomponents of BPI, we find that all subcomponents are negatively affected by mothers' participation in the program (depression, anxiousness and peerconflict are also significantly affected). More importantly, the subcomponents measuring anxiousness/depression and peerconflict show a significant increase. ITT and TOT show that the children of the treatment group mothers' have on an average 1.5-2% higher anxiousness and peerconflict than the control group mothers'. The program also decreases the PBI and all its subcomponents significantly. The PBI and all its subcomponents for the treatment group mothers' children are on an average 2-3% lower than those of the control group mothers'. The F statistic for the TOT is significant enough and does not point to weak instruments. These values are presented in column 5 of Table 1.6. The impact of the program on the mother though is positive. Assignment to the program increases the mothers' education level (significant at 5%) and also leads to higher earnings for the mother. It also leads to the mother providing a better home environment for the child as shown by the increase in the HOME score for mothers' who were in the treatment group. Since not all mothers' selected for the program participated fully in it, we try to measure the impact of an additional hour of the program on BBCS, BPI, PBI and the school performance.

$$Y_{ij} = \theta_o + \theta_1 \text{hours}_{ij} + \theta_2 X_{ij} + \delta_j + \varsigma_{ij} \tag{1.3.5}$$

where

$$\text{hours}_{ij} = \phi_o + \phi_1 Z_{ij} + \phi_2 X_{ij} + \delta_j + \xi_{ij} \tag{1.3.6}$$

Results are presented in Table 1.7. Since mothers' in the treatment group were supposed to attend the program for 20-30 hours per week for 18 months, we also calculate the total impact of the mother attending the program for an average of 25 hours per week in column 4 of Table 1.7. These results also indicate that mothers' participation in the program is decreasing the childs' cognitive score and increasing their behavioral problems.

## 1.4 Conclusions

We do not find any significant impact of investment in mothers' human capital have little impact on the cognitive and behavioral outcomes of their pre school children. The magnitudes of the effects through various channels like mothers' education, earnings and home score, are generally small, insignificant and are also likely to be of 'wrong' signs. However caution should be taken in generalizing them since these results are for a very select disadvantaged group of mothers' and seem to be driven by the problem of weak instruments. It is also possible that the effect of the program on the children was experienced with a lag and since there was no followup after 42 months we do not see any significant impact. Given the counterintutive signs we see on the impact of the mothers' variables on child outcomes, one area of further research this points to is the effects of reforms targeting mothers' self suffciency on the children.

## 1.5 Tables

Table 1.1: Summary Statistics

| Variable | Mean | Std Dev | Min | Max | N |
|---|---|---|---|---|---|
| Percent mothers' with higher than grade10 completed at baseline | 0.34 | 0.47 | 0 | 1 | 1842 |
| Test of Adult Basic Education score | 746.22 | 41.68 | 480 | 844 | 1836 |
| Age of mother at baseline | 18.82 | 1.35 | 16 | 22 | 1840 |
| Mothers' age at first birth | 16.82 | 1.37 | 13 | 19 | 1840 |
| Percent black | 0.55 | 0.5 | 0 | 1 | 1842 |
| Percent hispanic/others | 0.25 | 0.43 | 0 | 1 | 1842 |
| Dad with family at age 14 | 0.27 | 0.45 | 0 | 1 | 1842 |
| Percent mothers' with family never on welfare when young | 0.36 | 0.48 | 0 | 1 | 1825 |
| Age of child > 18 months at baseline | 0.43 | 0.5 | 0 | 1 | 1841 |
| Percent with male child | 0.52 | 0.5 | 0 | 1 | 1842 |
| At least one parent high school graduate or more | 0.49 | 0.5 | 0 | 1 | 1842 |
| Mothers' CESD depression score at baseline | 17.98 | 10.22 | 0 | 54 | 1838 |
| BBCS | 6.74 | 2.87 | 1 | 19 | 1754 |
| BPI | 109.43 | 13.42 | 68 | 145 | 1807 |

Table 1.2: Mean Comparison Tests

| Variable | Control | Treatment | Diff | p value | N_C | N_T |
|---|---|---|---|---|---|---|
| Percent mothers' with higher than grade10 completed at baseline | 0.33 | 0.35 | -0.02 | 0.45 | 609 | 1233 |
| Test of Adult Basic Education score | 745.48 | 746.58 | -1.09 | 0.6 | 607 | 1229 |
| Age of mother at baseline | 18.82 | 18.82 | -0.01 | 0.92 | 608 | 1232 |
| Percent black | 0.56 | 0.54 | 0.03 | 0.3 | 609 | 1233 |
| Percent hispanic/others | 0.24 | 0.26 | -0.02 | 0.47 | 609 | 1233 |
| Dad present at age 14 | 0.27 | 0.28 | -0.01 | 0.8 | 609 | 1233 |
| Percent mothers' with family never on welfare when young | 0.34 | 0.37 | -0.03 | 0.24 | 602 | 1223 |
| Age of child >18 months at baseline | 0.45 | 0.42 | 0.02 | 0.34 | 608 | 1233 |
| Percent with male child | 0.51 | 0.53 | -0.02 | 0.37 | 609 | 1233 |
| At least one parent high school graduate or more | 0.52 | 0.48 | 0.04 | 0.15 | 609 | 1233 |
| Mothers' CESD depression score at baseline | 18.38 | 17.78 | 0.6 | 0.24 | 607 | 1231 |

Table 1.3: OLS and IV: All Sample

| Variable | BBCS | | BPI | |
|---|---|---|---|---|
| | OLS | IV | OLS | IV |
| | (1) | (2) | (3) | (4) |
| Mother has high school/GED at month 42 | 0.477*** | -1.083 | -0.088 | 15.971 |
| | (0.154) | (1.876) | (0.746) | (11.228) |
| Mother has some college credits at month 42 | 0.694*** | 0.183 | -0.913 | 4.648 |
| | (0.216) | (1.709) | (1.035) | (9.837) |
| Home score at month 42 | 0.058*** | 0.103 | -0.242*** | 0.053 |
| | (0.007) | (0.090) | (0.033) | (0.488) |
| Mother earns $500 or more between months 31-42 | 0.113 | -0.115 | -0.249 | -12.304 |
| | (0.133) | (1.524) | (0.640) | (7.653) |
| Mother has higher than grade10 completed at baseline | 0.368** | 0.487 | 0.297 | -0.988 |
| | (0.148) | (0.314) | (0.710) | (1.654) |
| Mothers TABE score at baseline | 0.007*** | 0.010** | -0.009 | -0.044 |
| | (0.002) | (0.005) | (0.008) | (0.031) |
| Age of mother at baseline | 0.012 | 0.000 | 0.214 | 0.153 |
| | (0.054) | (0.062) | (0.256) | (0.359) |
| Black Dummy | -0.550*** | -0.472 | -3.526*** | -2.815 |
| | (0.211) | (0.444) | (1.025) | (2.230) |
| Hispanic Dummy | -0.674*** | -0.672*** | -2.633** | -2.231 |
| | (0.214) | (0.249) | (1.037) | (1.406) |
| Mothers' dad stayed with her at age 14 | -0.043 | -0.145 | -1.345* | -1.592 |
| | (0.152) | (0.200) | (0.736) | (1.056) |
| Mothers' family never on welfare when young | 0.223 | 0.274 | -0.235 | -1.284 |
| | (0.141) | (0.202) | (0.686) | (1.154) |

Standard error in parenthesis, covariates include family background and demographic variables

*p<0.1, **p<0.05, ***p<0.01

| Variable | BBCS | | BPI | |
|---|---|---|---|---|
| | OLS | IV | OLS | IV |
| | (1) | (2) | (3) | (4) |
| Mother ever married at baseline | -0.248 | -0.299 | 0.839 | 0.173 |
| | (0.237) | (0.269) | (1.143) | (1.543) |
| Age of child gt >18 months at baseline | 0.035 | -0.071 | -1.329** | -2.263 |
| | (0.139) | (0.334) | (0.677) | (1.873) |
| Male Child | -0.481*** | -0.427** | 0.481 | 0.455 |
| | (0.127) | (0.168) | (0.615) | (0.911) |
| At least one grandparent has completed high school | -0.079 | -0.073 | 0.807 | 0.156 |
| | (0.133) | (0.188) | (0.640) | (1.063) |
| Mothers' CESD depression score at baseline | -0.007 | -0.004 | 0.226*** | 0.254*** |
| | (0.006) | (0.010) | (0.031) | (0.053) |
| Constant | -4.702*** | -10.532 | 135.665*** | 130.981*** |
| | (1.721) | (6.792) | (8.258) | (34.703) |
| Site Fixed Effects | Yes | Yes | Yes | Yes |
| Observations | 1734 | 1734 | 1784 | 1784 |

Standard error in parenthesis, covariates include family background and demographic variables

*p<0.1, **p<0.05, ***p<0.01

Table 1.4: OLS and IV: Black, White and Hispanic

| Variable | BBCS | | BPI | |
|---|---|---|---|---|
| | OLS | IV | OLS | IV |
| | (1) | (2) | (3) | (4) |
| Panel A: Black | | | | |
| Mother has high school/GED at month 42 | 0.389* | -1.602 | -0.179 | 14.251 |
| | (0.217) | (1.842) | (1.003) | (9.087) |
| Mother has some college credits at month 42 | 0.386 | 1.014 | -0.972 | 3.229 |
| | (0.301) | (1.965) | (1.377) | (9.103) |
| Home score at month 42 | 0.075*** | 0.165** | -0.162*** | -0.232 |
| | (0.010) | (0.083) | (0.044) | (0.366) |
| Mother earns $500 or more between months 31-42 | 0.078 | 1.383 | -0.127 | -10.486 |
| | (0.188) | (1.541) | (0.865) | (6.983) |
| Observations | 938 | 938 | 968 | 968 |
| Panel B: White | | | | |
| Mother has high school/GED at month 42 | 0.750** | -5.391 | 1.138 | -37.853 |
| | (0.339) | (6.360) | (1.773) | (39.820) |
| Mother has some college credits at month 42 | 0.776 | 10.575 | -0.603 | 43.787 |
| | (0.526) | (8.626) | (2.735) | (49.748) |
| Home score at month 42 | 0.047*** | 0.059 | -0.394*** | 0.033 |
| | (0.015) | (0.205) | (0.079) | (1.117) |
| Mother earns $500 or more between months 31-42 | -0.168 | 3.068 | 0.030 | 4.553 |
| | (0.295) | (3.688) | (1.531) | (20.224) |
| Observations | 362 | 362 | 363 | 363 |
| Panel C: Hispanic | | | | |
| Mother has high school/GED at month 42 | 0.224 | -4.049 | -1.285 | 12.537 |
| | (0.300) | (4.443) | (1.661) | (25.698) |
| Mother has some college credits at month 42 | 1.078*** | -0.381 | -1.643 | -12.049 |
| | (0.392) | (4.181) | (2.181) | (19.742) |
| Home score at month 42 | 0.020 | 0.242 | -0.275*** | -0.787 |
| | (0.013) | (0.260) | (0.074) | (1.082) |
| Mother earns $500 or more between months 31-42 | 0.536** | -1.516 | -0.413 | 15.641 |
| | (0.253) | (2.858) | (1.396) | (15.469) |
| Observations | 402 | 402 | 415 | 415 |

All regressions include covariates and site fixed effects

Standard error in parenthesis, covariates include family background and demographic variables

*$p<0.1$, **$p<0.05$, ***$p<0.01$

Table 1.5: OLS and IV: Child's age gt/lt 18 months

| Variable | BBCS | | BPI | |
|---|---|---|---|---|
| | OLS | IV | OLS | IV |
| | (1) | (2) | (3) | (4) |
| Panel A: Focal Child $\geq$ 18 months | | | | |
| Mother has high school/GED at month 42 | 0.081 | 2.390 | -1.380 | 4.044 |
| | (0.251) | (2.620) | (1.166) | (12.171) |
| Mother has some college credits at month 42 | 0.259 | 1.296 | -1.201 | 7.427 |
| | (0.343) | (2.461) | (1.557) | (10.192) |
| Home score at month 42 | 0.057*** | 0.121 | -0.275*** | -0.042 |
| | (0.011) | (0.076) | (0.051) | (0.344) |
| Mother earns $500 or more between months 31-42 | 0.229 | -1.379 | -1.077 | -6.523 |
| | (0.209) | (2.264) | (0.968) | (8.895) |
| Observations | 716 | 716 | 767 | 767 |
| Panel A: Focal Child < 18 months | | | | |
| Mother has high school/GED at month 42 | 0.639*** | -3.061 | 0.958 | 18.218* |
| | (0.194) | (2.316) | (0.987) | (10.806) |
| Mother has some college credits at month 42 | 0.997*** | -0.771 | -0.543 | 1.836 |
| | (0.279) | (3.359) | (1.408) | (15.536) |
| Home score at month 42 | 0.057*** | 0.201 | -0.217*** | -0.135 |
| | (0.009) | (0.129) | (0.044) | (0.599) |
| Mother earns $500 or more between months 31-42 | 0.151 | 0.975 | 0.268 | -9.483 |
| | (0.172) | (1.754) | (0.870) | (7.546) |
| Observations | 1018 | 1018 | 1017 | 1017 |

All regressions include covariates and site fixed effects

Standard error in parenthesis, covariates include family background and demographic variables

*p<0.1, **p<0.05, ***p<0.01

Table 1.6: Intent to Treat and Treatment on the Treated

| Variable | Control group mean | ITT | Naive TOT | TOT | F Statistic | N |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| BBCS | 6.836 | -0.206 | -0.147 | -0.231 | 5008.192 | 1,734 |
| | | (0.139) | (0.135) | (0.154) | | |
| BPI | 108.489 | 1.562** | 1.418** | 1.752** | 5123.908 | 1,784 |
| | | (0.661) | (0.639) | (0.735) | | |
| BPI-Anxious/ Depressed | 105.722 | 1.694*** | 1.327** | 1.897*** | 5243.647 | 1,805 |
| | | (0.579) | (0.561) | (0.644) | | |
| BPI-Antisocial | 109.978 | 0.525 | 0.759 | 0.589 | 5149.509 | 1,788 |
| | | (0.717) | (0.694) | (0.797) | | |
| BPI-Dependent | 108.156 | 0.956 | 0.787 | 1.071 | 5235.125 | 1,804 |
| | | (0.629) | (0.609) | (0.699) | | |
| BPI-Headstrong | 102.317 | 0.368 | 0.596 | 0.413 | 5196.83 | 1,806 |
| | | (0.580) | (0.561) | (0.645) | | |
| BPI-Hyperactive | 107.826 | 1.530** | 1.234* | 1.714** | 5205.323 | 1,807 |
| | | (0.670) | (0.649) | (0.745) | | |
| BPI-Peerconflict | 105.982 | 2.021*** | 1.912*** | 2.264*** | 5240.513 | 1,804 |
| | | (0.670) | (0.649) | (0.745) | | |
| PBI | 197.382 | -4.563*** | -3.919** | -5.114*** | 5177.658 | 1,792 |
| | | (1.635) | (1.584) | (1.819) | | |
| PBI-Autonomous | 43.394 | -0.905*** | -0.596* | -1.015*** | 5216.33 | 1,800 |
| | | (0.335) | (0.325) | (0.373) | | |
| PBI-Compliant | 63.259 | -2.165*** | -1.868** | -2.427*** | 5177.658 | 1,792 |
| | | (0.765) | (0.741) | (0.851) | | |
| PBI-Sensitive | 90.71 | -1.469** | -1.427** | -1.646** | 5186.636 | 1,794 |
| | | (0.724) | (0.701) | (0.805) | | |
| School Performance | 8.341 | -0.127 | -0.079 | -0.144 | 1937.192 | 780 |
| | | (0.152) | (0.146) | (0.169) | | |
| Mother has high school/GED at month 42 | | 0.038* | | | | 1,819 |
| | | (0.023) | | | | |
| Mother has some college credits at month 42 | | 0.036** | | | | 1,819 |
| | | (0.016) | | | | |
| Mother earns $500 or more between months 31-42 | | 0.007 | | | | 1,819 |
| | | (0.024) | | | | |
| Home score at month 42 | | 0.287 | | | | 1,761 |
| | | (0.492) | | | | |

All regressions include covariates and site fixed effects

Standard error in parenthesis, covariates include family background and demographic variables

*p<0.1, **p<0.05, ***p<0.01

Table 1.7: IV: Number of New Chance Hours

| Variable | Mean of group with zero hours | OLS Estimates | IV of hours at the program | 1800 hours | F Statistic | N |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| BBCS | 6.727 | 0.0005** | -0.001 | -1.8 | 547.635 | 1734 |
| | | (0.000) | (0.000) | | | |
| BPI | 108.647 | 0.001 | 0.005** | 9 | 572.265 | 1784 |
| | | (0.001) | (0.002) | | | |
| BPI-Anxious/ Depressed | 106.019 | -0.000 | 0.005*** | 9 | 577.775 | 1805 |
| | | (0.001) | (0.002) | | | |
| BPI-Antisocial | 109.831 | 0.001 | 0.002 | 3.6 | 575.002 | 1788 |
| | | (0.001) | (0.002) | | | |
| BPI-Dependent | 108.37 | 0.0003 | 0.003 | 5.4 | 575.043 | 1804 |
| | | (0.001) | (0.002) | | | |
| BPI-Headstrong | 102.127 | 0.0002 | 0.001 | 1.8 | 575.164 | 1806 |
| | | (0.001) | (0.002) | | | |
| BPI-Hyperactive | 108.068 | 0.001 | 0.005** | 9 | 576.619 | 1807 |
| | | (0.001) | (0.002) | | | |
| BPI-Peerconflict | 106.279 | 0.001 | 0.006*** | 10.8 | 577.549 | 1804 |
| | | (0.001) | (0.002) | | | |
| PBI | 196.556 | -0.0002 | -0.014*** | -25.2 | 571.468 | 1792 |
| | | (0.003) | (0.005) | | | |
| PBI-Autonomous | 43.129 | 0.0003 | -0.003*** | -5.4 | 572.618 | 1800 |
| | | (0.001) | (0.001) | | | |
| PBI-Compliant | 62.891 | -0.001 | -0.007*** | -12.6 | 571.468 | 1792 |
| | | (0.001) | (0.002) | | | |
| PBI-Sensitive | 90.522 | 0.000 | -0.005** | -9 | 569.956 | 1794 |
| | | (0.001) | (0.002) | | | |
| School Performance | 8.294 | 0.0001 | -0.0004 | -0.72 | 269.528 | 780 |
| | | (0.000) | (0.000) | | | |

All regressions include covariates and site fixed effects

Standard error in parenthesis, covariates include family background and demographic variables

*p<0.1, **p<0.05, ***p<0.01

Table 1.8: Distribution across sites

| Site | Control | Treatment | Total |
|---|---|---|---|
| Allentown | 35 | 67 | 102 |
| Bronx | 39 | 84 | 123 |
| Chicago | 16 | 38 | 54 |
| Chulavista | 38 | 71 | 109 |
| Denver | 32 | 63 | 95 |
| Detroit | 53 | 101 | 154 |
| Harlem | 38 | 73 | 111 |
| Inglewood | 41 | 78 | 119 |
| Jacksonville | 40 | 88 | 128 |
| Lexington | 43 | 73 | 116 |
| Minneapolis | 34 | 70 | 104 |
| Philadelphia | 44 | 85 | 129 |
| Pittsburgh | 48 | 103 | 151 |
| Portland | 38 | 87 | 125 |
| Salem | 31 | 69 | 100 |
| San Jose | 39 | 83 | 122 |

Table 1.9: First Stage F Statistics

| Variable | Mother has high school/GED at month 42 | Mother has some college credits at month 42 | Home score at month 42 | Mother earns > 500 or more between months 31-42 |
|---|---|---|---|---|
| BBCS all | 0.37 | 1.3 | 0.48 | 0.95 |
| | (0.96) | (0.23) | (0.91) | (0.49) |
| BPI all | 0.39 | 1.16 | 0.37 | 0.94 |
| | (0.95) | (0.32) | (0.96) | (0.5) |
| Focal Child Age gt 18 months BBCS | 0.54 | 0.64 | 0.62 | 1.3 |
| | (0.9) | (0.82) | (0.84) | (0.21) |
| Focal Child Age gt 18 months BPI | 0.61 | 0.63 | 0.65 | 1.22 |
| | (0.84) | (0.82) | (0.81) | (0.26) |
| Focal Child Age lt 18 months BBCS | 1.19 | 1.89 | 1.16 | 1.74 |
| | (0.27) | (0.02) | (0.29) | (0.04) |
| Focal Child Age lt 18 months BPI | 1.35 | 1.95 | 1.06 | 1.6 |
| | (0.17) | (0.02) | (0.39) | (0.07) |
| Black BBCS | 0.74 | 0.86 | 0.63 | 1.39 |
| | (0.76) | (0.62) | (0.86) | (0.14) |
| Black BPI | 0.8 | 0.85 | 0.63 | 1.3 |
| | (0.69) | (0.62) | (0.86) | (0.19) |
| White BBCS | 0.61 | 1.79 | 1.37 | 1.38 |
| | (0.88) | (0.03) | (0.15) | (0.14) |
| White BPI | 0.74 | 1.65 | 1.33 | 1.18 |
| | (0.75) | (0.05) | (0.17) | (0.28) |
| Hispanic BBCS | 0.9 | 1.65 | 0.75 | 1.7 |
| | (0.57) | (0.05) | (0.74) | (0.04) |
| Hispanic BPI | 1.14 | 1.61 | 0.66 | 1.53 |
| | (0.31) | (0.06) | (0.83) | (0.08) |

p values in parenthesis

# Bibliography

**Auld, M. Christopher, and Nirmal Sidhu.** 2005. "Schooling, Cognitive Ability and Health." *Health Economics*, 14(10): 1019–1034.

**Bloom, Howard S.** 2005. *Learning More from Social Experiments: Evolving Analytic Approaches.*: NY: Russell Sage Foundation, 246 pp.

**Bowles, Samuel, Herbert Gintis, and Melissa Osborne.** 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature*, 39(4): 1137–1176.

**Brooks-Gunn, Jeanne, Wen-Jui Han, and Jane Waldfogel.** 2002. "Maternal Employment and Child Cognitive Outcomes in the First Three Years of Life: The NICHD Study of Early Child Care." *Child Development*, 73(4): 1052–1072.

**Carlson, Marcia J., and Mary E. Corcoran.** 2001. "Family Structure and Children's Behavioral and Cognitive Outcomes." *Journal of Marriage and the Family*, 63(3): 779–792.

**Cawley, John, James J. Heckman, and Edward Vytlacil.** 2001. "Three Observations on Wages and Measured Cognitive Ability." *Labour Economics*, 8(4): 419–442.

**Cunha, Flavio, and James J. Heckman.** 2008. "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Journal of Human Resources*, 43(4): .

**Desai, Sonalde, P. Lindsay Chase-Lansdale, and Robert T. Michael.** 1989. "Mother or Market? Effects of Maternal Employment on the Intellectual Ability of 4-Year-Old Children." *Demography*, 26(4): 545–561.

**Heckman, James J., and Yona Rubinstein.** 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *The American Economic Review*, 91(2): 145–149.

**Heckman, James J., Jora Stixrud, and Sergio Urzua.** 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3): 411–482.

**Herrnstein, Richard J., and Charles A. Murray.** 1994. *The Bell Curve : Intelligence and Class Structure in American Life.*: Free Press, New York :, xxvi, 845 p. :.

**Krueger, Alan B.** 2003. "Economic Considerations and Class Size." *The Economic Journal*, 113(485): F34–F63.

**Leibowitz, Arleen.** 1974. "Education and Home Production." *The American Economic Review*, 64(2): 243–250.

**Magnuson, Katherine.** 2007. "Maternal Education and Children's Academic Achievement During Middle Childhood." *Developmental Psychology*, 43.

**Murnane, Richard J., Rebecca A. Maynard, and James C. Ohls.** 1981. "Home Resources and Children's Achievement." *The Review of Economics and Statistics*, 63(3): 369–377.

**Murnane, Richard J., John B. Willett, and Frank Levy.** 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *The Review of Economics and Statistics*, 77(2): 251–266.

**Neal, Derek A., and William R. Johnson.** 1996. "The Role of Premarket Factors in Black-White Wage Differences." *The Journal of Political Economy*, 104(5): 869–895.

**Rosenzweig, Mark R., and Kenneth I. Wolpin.** 1994. "Are There Increasing Returns to the Intergenerational Production of Human Capital? Maternal Schooling and Child Intellectual Achievement." *The Journal of Human Resources*, 29(2): 670–693.

**Todd, Petra E., and Kenneth I. Wolpin.** 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *The Economic Journal*, 113(485): 3–33.

**Todd, Petra E., and Kenneth I. Wolpin.** "The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps." latest version, available at `http://www.ssc.upenn.edu/ier/Archive/04-019.pdf`.

Chapter 2

# Personality and Entrepreneurship

## 2.1 Introduction

Entrepreneurship is an important area of research in economics. According to Hipple (2010) : In 2009, 15.3 million individuals were self-employed, including both those who had incorporated their businesses and those who did not. The self-employment rate, which is the proportion of total employment made up of the self-employed, was 10.9 percent. Entrepreneurship is seen as essential for an economy to grow and flourish. Entrepreneurial activity is usually seen as integral step toward innovation and globalization. According to UNCTAD: *Entrepreneurs produce solutions that fly in the face of established knowledge, and they always challenge the status quo. They are risk-takers who pursue opportunities that others may fail to recognize or may even view as problems or threats. Whatever the definition of entrepreneurship, it is closely associated with change, creativity, knowledge, innovation and flexibility-factors that are increasingly important sources of competitiveness in an increasingly globalized world economy. Thus, fostering entrepreneurship means promoting the competitiveness of businesses.*

Economists have long examined various aspects of entrepreneurship. Using data from the Panel Study of Income Dynamics, Quadrini (1999) finds that there is a marked concentration of wealth in the hands of entrepreneurs and that entrepreneurs experience greater upward mobility than workers. Simultaneously, researchers claim that non-availability of the required funds is one of the main constraints that the potential entrepreneurs face(Evans and Jovanovic 1989; Holtz-Eakin et al. 1994). Wealthier households are more likely to start a business. Paulson and Townsend (2004) conclude that financial constraints serve an important role in determining the shape of the patterns of entrepreneurship in Thailand. Paulson et al. (2006) estimate a model in which the choice between entrepreneurship and wage work may be influenced by financial market imperfections. They conclude that moral hazard is the key financial constraint that restricts entrepreneurship in Thailand. Literature examining the earnings differential between self-employed and paid workers points out that even though paid workers on an average earn more than self-employed people, self-employed have greater job satisfaction. Evans and Leighton (1989) conclude that their results are consistent with the disadvantage theory which views entrepreneurs as misfits cast off from wage work. According to them people who switch from wage work to self employment tend to be people who were receiving relatively low wages, who have changed jobs frequently, and who experience relatively frequent or long spells of unemployment as wage workers. Hamilton (2000) examines reasons for earnings differentials between paid employment and entrepreneurship. He finds that entrepreneurs have both lower initial earnings and lower earnings growth than in paid employment. His conclusion is that the differential cannot be explained by the selection of low-ability employees into self-employment, instead the self-employment earnings differential reflects entrepreneurs' willingness to sacrifice substantial earnings in exchange for the non-pecuniary benefits of owning a business. Blanchflower and Oswald (1998) conclude that the self-employed have higher levels of job and life satisfaction than employees. Most of this literature concludes that self-employed people earn less than wage workers but have higher levels of job satisfaction.

One of the reasons for entrepreneurship has been called the intergenerational pick up rate with respect to self employment by Hout and Rosen (2000). They conclude that the primary factor affecting an individuals self employment is the self employment status of his or her father. The nature vs. nurture debate is an ongoing one in entrepreneurship. Thus family background (nurture) has been examined as a reason for self employment. More recently researchers have ventured into the field of behavioral genetics (nature) as a possible explanation for people venturing into entrepreneurship. The ACE model that divides an observed trait into a genetic component (A), a shared environmental component (C), and a unique environmental component (E), has been used here. Recently, Nicolaou and colleagues (Nicolaou & Shane,2009; Nicolaou, Shane, Cherkas, Hunkin, & Spector, 2008; Nicolaou,Shane, Cherkas, & Spector, 2008 and Zhang et al. 2009) have conceptually argued, and have provided empirical evidence for a genetic underpinning of entrepreneurship. Thus, the question examining why people choose self-employment is central to economic research on entrepreneurship. There is not yet a consensus on this question. Researchers have examined several possible explanations including being your own boss (Hamilton 2000), race (Fairlie and Robb 2007) and gender (Devine 1994). The relation between cognitive ability and labor market outcomes has long been studied (Boissiere et al. 1985; Cawley et al. 2001; Murnane et al. 1995). More recently the importance of non-cognitive ability on labor market outcomes has also been documented (Bowles et al. 2001; Goldsmith et al. 1997; Heckman et al. 2006). This paper attempts to examine the relation between non cognitive abilities as measured by the big five personality traits and self employment.

Most of the existing research on personality and occupation has been on the lines of either personality and preferences or personality and performance. The personality and preference literature looks at why different people enter different occupations. This literature studies the relationship between individual tastes and preferences and the occupation in which an individual finds employment. Hence an individual chooses to go into an occupation that provides her satisfaction. One of the earlier papers in this literature is by Filer (1986) who uses personality and tastes to predict which of five broadly defined occupational groups an individual will enter. Antecol and Cobb-Clark (2010) investigate the role of non-cognitive skills in the occupational segregation of young workers entering the U.S. labor market. They find entry into male-dominated fields of study and male-dominated occupations are both related to the extent to which individuals believe they are intelligent and have male traits while entry into male-dominated occupations is also related to the willingness to work hard, impulsivity, and the tendency to avoid problems. Using an assignment model Borghans et al. (2006) show that people are most productive in jobs that match their style and earn less when they have to shift to other jobs. Krueger and Schkade (2008) show that workers who are more gregarious, based on their behavior off the job, tend to be employed in jobs that involve more social interactions. The literature on personality and performance works on the assumption that people choose jobs where the returns to their personality type would be the highest. In this literature a job is usually a source of income and people tend to go for go for jobs that provide the highest compensation to their personality type. Fortin (2008) examines the impact of four non-cognitive traits: self-esteem, external locus of control, the importance of money/work and the importance of people/ family on wages and on the gender wage gap. He finds that gender differences in these non-cognitive factors, especially the importance of money/work, have a modest but significant role in accounting for the gender wage gap. Urzua (2008) finds that the effects of non cognitive ability on schooling decisions, hourly wages and annual hours worked are are uniformly stronger for blacks than whites. In this paper we try to bring these two literatures together by estimating the impact of personality traits on the choice to be self employed as well as on the entrepreneurship income.

The rest of the paper proceeds as follows: section 2 is a literature survey on the big five personality traits and their importance in the psychology and more recently in the economics literature, section 3 presents the model along with the estimation strategy, the dataset used along with summary statistics is presented in section 4, results are in section 5 and section 6 concludes.

## 2.2 Background Literature

The importance of personality traits has been revealed by the inability of cognitive ability to predict certain outcomes. Heckman and Rubinstein (2001) use evidence from the General Education Development (GED) testing program to demonstrate the importance of personality traits. The level of cognitive ability of GED recipients is the same as high school graduates who do not go on to college as measured by scores on the Armed Forces Qualifying Test (AFQT). Controlling for cognitive ability, GED recipients have lower hourly wages than high school dropouts. GED recipients also have higher job turnover rates, and are more likely to drop out of the army and post secondary schooling (Heckman and LaFontaine 2010). Heckman, Stixrud, and Urzua (2006) find that the power of the personality traits equals or exceeds the predictive power of cognitive traits for schooling, occupational choice, wages, health behaviors, teenage pregnancy and crime. Additionally Heckman, Stixrud, and Urzua (2006) and Judge and Hurst (2007) show that among participants in the NLSY 1979 cohort, positive self-evaluations measured in young adulthood (with positive self-evaluations of self-esteem, locus of control, and related traits) predict income in mid-life and, further, enhance the benefits of family socioeconomic status, and academic achievement on mid-life income.

Personality researchers have proposed that there are five basic dimensions of personality. These five factors have been known as the Big Five since Goldberg (1971). Based on the research by Goldberg (1981, 1993) and McCrae and Costa (1987, 1997), these five categories are usually described as follows:

1. Extraversion: This trait includes characteristics such as excitability, sociability, talkativeness, assertiveness, and high amounts of emotional expressiveness.

2. Agreeableness: This personality dimension includes attributes such as trust, altruism, kindness, affection, and other pro-social behaviors.

3. Conscientiousness: Common features of this dimension include high levels of thoughtfulness, with good impulse control and goal-directed behaviors. Those high in conscientiousness tend to be organized and mindful of details.

4. Neuroticism: Individuals high in this trait tend to experience emotional in-

stability, anxiety, moodiness, irritability, and sadness.

5. Openness: This trait features characteristics such as imagination and insight, and those high in this trait also tend to have a broad range of interests.

These factors represent personality at a broad level of aggregation. This categorization does not imply that all personality attributes can be fully reduced to five traits. Each factor summarizes a large number of distinct, more specific, personality characteristics. To provide a better idea of what they are, in Table 2.1 we list a number of characteristics related to each of the five personality dimensions. Research demonstrates that these factors and facets are generally stable across the lifespan (Roberts & DelVecchio 2000). According to Sutin et al. (2009) when measured concurrently, controlling for sex, ethnicity, age and education, personality was associated with income and job satisfaction: Emotionally stable and conscientious participants reported earning higher incomes and reported more satisfaction with their jobs. The Big Five have been used extensively in psychology literature to predict labor market and social outcomes (Barrick et al. 1993; Groves 2005; Kanfer et al. 2001). Judge et al. (1999) report a near consensus in the organizational psychology literature that out of these five traits conscientiousness, extraversion, and neuroticism are most relevant to job performance. Psychology literature also has substantial evidence on the importance of personality traits in predicting socioeconomic outcomes including job performance, health, and academic achievement (Barrick and Mount 1991; Chamorro-Premuzic and Furnham 2005; Hampson et al. 2006; Hogan, Hogan, and Roberts 1996; Hogan and Holland 2003; Robbins et al. 2006; Roberts et al. 2007, Ones et al. 2007; Schmidt and Hunter 1998).

More recently economists have also started using these traits to predict various outcomes. Heineck (2011) analyses British Household Panel Survey data using the five factor model to examine the relationship between individuals' personality and wages in the UK. He finds a negative linear relationship between wages and agreeableness and, for females, wages and neuroticism whereas openness to experience is rewarded. Anger and Heineck (2010) do a joint analysis of the relationship between cognitive skills, personality traits and earnings in Germany. They find that personality is an important predictor of earnings even if a large set of socio-demographic and job-related characteristics and, even more relevant, cognitive ability scores are included. Mueller and Plug (2006) use the Five-Factor Model of personality structure to explore how personality affects the earnings of a large group of men and women. They find that all five basic traits: extroversion, agreeableness, conscientiousness, neuroticism and openness to experience had statistically significant positive or negative earnings effects, and together they appear to have had effects comparable to those commonly found for cognitive ability. Using data from the military enlistment for a large representative sample of Swedish men, Lindqvist and Vestman (2011) find strong evidence that men who fare badly in the labor market in the sense of long-term unemployment or low annual earnings lack non-cognitive but not cognitive ability.

However, still the links between measures of personality and preferences are largely unexplored. Do preferences causally effect personality? Does personality causally effect preferences? Or are both effected from other parameters? Some aspects of personality may be reflecting preferences. For example, Openness to Experience might relate to a preference for learning, and Extraversion might reflect a preference for social interactions. There have been few studies that empirically investigate the link between preference and personality. The most examined facets of preferences are time preference, risk aversion and leisure. Daly, Delaney and Harmon (2009) find that financial discounting is related to a range of psychological variables including consideration of future consequences, self-control,conscientiousness, extraversion, and experiential avoidance. Borghans, Meijers and ter Weel (2008) examine whether non-cognitive skills, measured both by personality traits and by economic preference parameters, influence cognitive tests' performance. Their basic idea is that non-cognitive skills might affect the effort people put into a test to obtain good results. Dohmen, Falk, Huffman et al. (2010) conclude that Openness to Experience and Agreeableness are related to risk aversion. Borghans, Golsteyn, Heckman et al. (2009) show that risk-aversion is positively associated with Neuroticism, which contains measures of fear and strong emotional responses to bad outcomes. They also find that Agreeableness is positively associated with risk aversion. Anderson, Burks, DeYoung et al. (2011) find that Neuroticism is positively associated with risk aversion but only for lotteries over gains not losses.(Find Citation) Barsky, Juster, Kimball et al. (1997) measure risk tolerance, time preference, and the intertemporal elasticity of substitution and find that risk tolerance predicts smoking and drinking, holding insurance and stock, and decisions to immigrate and be self-employed.

The relation between personality and entrepreneurship has also been examined both in the psychology and economics literature. We would expect individuals to be attracted to entrepreneurship based on the self-perceived match between their own personality traits and the task demands of entrepreneurship. By the same logic, we expect people who have more of the personality traits associated with the entrepreneurial role to be more successful entrepreneurs. Using meta-analytical techniques to examine the relationship between personality and entrepreneurial status, Zhao and Seibert (2006) conclude that entrepreneurs are higher on conscientiousness, emotional stability, and openness to experience and are lower on agreeableness than non-entrepreneur managers. In a more recent study Zhao et al. (2010) find that openness to experience and conscientiousness appear to be the personality constructs most strongly and consistently associated to both entrepreneurial intentions and entrepreneurial performance. They claim that personality plays a role both in the intention to become an entrepreneur and success as an entrepreneur. On similar lines Rauch and Frese (2007) find that the traits matched to entrepreneurship significantly correlated with entrepreneurial behavior (business creation, business success) were need for achievement, generalized self-efficacy, innovativeness, stress tolerance, need for autonomy, and proactive personality.

## 2.3 Empirical Framework

In this section we develop the empirical framework relating personality traits to performance and preferences for entrepreneurship. Our approach is based on the Roy (1951) model of occupational choice in which individuals sort themselves across sectors based on relative abilities. We follow Heckman, Stixrud, and Urzua (2006) who extend the Roy model to incorporate psychological variables that affect occupational choice and performance.[1] Individuals will choose entrepreneurship in a given

[1] Almlund et al (2011) develop a generalized version of the Roy model incorporating personality traits, multiple tasks, goals, and effort. Their model informs the approach taken here.

period if the expected utility from self-employment is greater than that of wage work. Since our data sample will consist of prime-age males, we do not consider a non-employment option, although this could easily be introduced. Utility in each sector is a function of income as well as preferences for non-wage attributes associated with the sector. For example, these preferences would include the desire to "be your own boss" in self-employment. We write the utility from self-employment, $V_{iSE}$, as:

$$V_{iSE} = \alpha y_{iSE} + \delta_{SE} P_i + Z_i \gamma_{SE} + \eta_{iSE} \tag{2.3.1}$$

where $y_{iSE}$ is a measure of income if the individuals is self-employed in the current period, $P_i$ is a vector of the personality traits of individual $i$ that affect the utility associated with entrepreneurship, $Z_i$ is a vector of other observed individual characteristics that affect utility, such as education, and $\eta_{iSE}$ are unobserved (to the econometrician) factors that influence utility. We specify a similar equation for the utility associated with paid employment:

$$V_{iPE} = \alpha y_{iPE} + \delta_{PE} P_i + Z_i \gamma_{PE} + \eta_{iPE} \tag{2.3.2}$$

At the start of the period, an individual will choose to be an entrepreneur if the expected utility from self-employment is greater than that from paid employment; he will choose paid employment if the opposite is true. Let the index $I_i^*$ be the difference in the expected utilities across sectors:

$$
\begin{aligned}
I_i^* &= E[V_{iSE}] - E[V_{iPE}] \\
&= \alpha(E[y_{iSE}] - E[y_{iPE}]) + (\delta_{SE} - \delta_{PE})P_i + Z_i(\gamma_{SE} - \gamma_{PE}) + (\eta_{iSE} - \eta_{iPE})
\end{aligned}
\tag{2.3.3}
$$

Equation (2.3.3) implies that individuals do not know their self-employment (or paid employment) income when making their employment decision, but do know their preferences for the non-wage attributes of each sector. Individuals are thus observed to be self-employed, denoted by $I_i = 1$, if $I_i^* > 0$, and are paid employees if $I_i^* \leq 0$ $(I_i = 0)$. Since choice is based on relative utility, we are only able to recover differences in the parameters in equation (2.3.3), which we rewrite as:

$$
\begin{aligned}
I_i^* &= E[V_{iSE}] - E[V_{iPE}] \\
&= \alpha(E[y_{iSE}] - E[y_{iPE}]) + \delta P_i + Z_i\gamma + \eta_i
\end{aligned}
\tag{2.3.4}
$$

Turning to performance, we specify the sectoral annual income equations as:

$$
y_{iSE} = X_i\beta_{SE} + \pi_{SE}P_i + \varepsilon_{iSE} \tag{2.3.5}
$$

and

$$
y_{iPE} = X_i\beta_{PE} + \pi_{PE}P_i + \varepsilon_{iPE} \tag{2.3.6}
$$

The vector $X_i$ include factors thought to influence productivity and income, such as experience. The error terms $\varepsilon_{iSE}$ and $\varepsilon_{iPE}$ represent shocks to sectoral returns that are not known by individual $i$ when he chooses whether to be an entrepreneur at the start of the period. The parameter vectors $\pi_{SE}$ and $\pi_{PE}$ are of particular interest since they measure the impact of personality traits in self-employment and paid-employment, respectively.

The goal of this study is to distinguish the impact of personality traits on labor market performance, as given by $\pi_{SE}$ and $\pi_{PE}$, from the relationship between personality and preferences for entrepreneurship, as given by $\delta$. Prior studies of the relationship between personality and occupational choice generally do not estimate a structural probit model like equation (2.3.4) that explicitly incorporates expected earnings. Instead, studies typically estimate a reduced-form probit model that substitutes the expected values of equations (2.3.5) and (2.3.6) to generate a model of the form:

$$\begin{aligned} I_i^* &= \alpha\big((X_i\beta_{SE} + \pi_{SE}P_i) - (X_i\beta_{PE} + \pi_{PE}P_i)\big) + \delta P_i + Z_i\gamma + \eta_i \\ &= (\alpha\pi_{SE} - \alpha\pi_{PE} + \delta)P_i + X_i(\alpha\beta_{SE} - \alpha\beta_{SE}) + Z_i\gamma + \eta_i \\ &= \tau P_i + X_i\kappa + Z_i\gamma + \eta_i \end{aligned} \tag{2.3.7}$$

As is clear from equation (2.3.7), the coefficient vector $\tau$ associated with $P_i$ recovered from the reduced-form probit model reflects both the performance and preference implications of personality traits. For example, suppose that estimation of the reduced-form probit model (2.3.7) shows that individuals with higher levels of conscientiousness are more likely to be entrepreneurs. It is unclear whether this reflects a greater impact of conscientiousness on entrepreneurial performance (e.g., $\pi_{SE} > \pi_{PE}$), or stronger preference for the non-pecuniary aspects of self-employment among these individuals (e.g., $\delta > 0$), or both. By focusing on the estimation of equation (2.3.4), we will be able to distinguish between these alternative explanations.

Estimation of equation (2.3.4) requires specification of individual $i$'s expected return in each sector, $E[y_{iSE}]$ and $E[y_{iPE}]$. Since individuals may non-randomly sort into paid employment and self-employment, simple OLS estimation of the sectoral income equations (2.3.5) and (2.3.6) may yield biased estimates of the parameters $\pi_{SE}$, $\pi_{PE}$, $\beta_{SE}$, and $\beta_{PE}$ if the unobserved (by the researcher) factors influencing income, $\varepsilon_{iSE}$ and $\varepsilon_{iPE}$, are correlated with the unobserved preference component $\eta_i$. Consequently, we estimate equations (2.3.4) and (2.3.5) incorporating inverse mills-ratios generated from equation (2.3.6) to account for potential self-selection, which allows us to recover unbiased estimates of sectoral income equation parameters. These parameters will be used to construct the individual's expected difference in sectoral returns, $E[y_{iSE}] - E[y_{iPE}]$, which is then included in the structural probit model (2.3.4) as a regressor.

## 2.4 Data

The data for this paper is from The Midlife Development in the United States (MIDUS) carried out in 1994/95 by the MacArthur Midlife Research Network. The first wave of data collection (MIDUS I) began in 1995 and did a national survey of over $7,000$ Americans aged $25$ to $74$. The main data collection consisted of a general population survey, as well as surveys of siblings of the general population respondents, and a twin pairs sample. The MIDUS II project was designed to collect a second wave of data on the same respondents approximately $10$ years later. In addition to the national probability sample (N = $3485$), the study included over-samples in select metropolitan areas (N = $757$), a sample of siblings (N = $951$) of the main respondents, and a national sample of twin pairs (N = $1914$). The purpose of the study was to investigate the role of behavioral, psychological, and social factors in understanding age-related differences in physical and mental health. The study collected extensive information on the personality traits as well as the socioeconomic characteristics of the sample. We use this data from both the waves to examine the relation between the personality traits and self employment decision as well as success. The analysis was done on the male working subpopulation aged 65 or less to avoid modelling the decision to work for women. Since data for only a subsample of the people interviewed in 1995 was collected in 2004, our final sample is for 1100 males.

Table 2.2 presents the summary statistics of the sample for 2004. In Table 2.3, we present the t-tests to see the difference in the type of people choosing self employment vs. paid employment. The MIDUS male working sample is around 50 years of age in 2004 with an annual income of $\$70,000$. $79\%$ of them are married and the average male has 3 children in 2004. In terms of education 4% of them have a GED or lower, 20% are high school graduates, 26% have some college and 49% have a graduate degree or higher. In 2004 22% of the sample is self employed. We did some t-tests for the difference between self employed and paid employees for 2004. Self employed people are significantly older than paid employees. Education wise there is no significant difference between the two groups in 2004. The self employed have a higher annual income while there is no significant difference between the proportion of people married and the number of children. Results are presented in Table 2.3. In the same table we also present the t-tests for the difference in personality between self employed and paid employees in our sample. Self employed people score higher on the extraversion and openness to experience.

## 2.5 Results

We begin our analysis by estimating the reduced-form probit model of self-employment choice in 2004 given by equation (2.3.7). The model includes the standard set of demographic variables, such as age and education, the Big 5 personality measures, and a rich set of family background variables that are available in the MIDUS data, including father's and mother's education, whether the father and mother were present in the home when the individual was 14 years old, and whether the father and mother had been self-employed when the individual was growing up. Similar to prior studies incorporating personality characteristics, Table 2.4 shows that individuals who are more open to new experiences are significantly more likely to be entrepreneurs in 2004. While agreeableness and neuroticism have a negative effect on self-employment, these variables are not statistically significant. Somewhat surprisingly, education is not a significant predictor of self-employment, nor is age or marital status. One might have expected that more educated or older individuals would have more access to credit which would increase the probability of entrepreneurial entry (Evans and Jovanovic 1989). A notable finding is the asymmetry in the roles that fathers and mothers play in the likelihood of entrepreneurship. Having a self-employed father is a strong positive predictor of entrepreneurial choice, even though the individuals in our sample are on average 50 years old, implying that the father's occupational status 35 years ago still affects choice today. Somewhat surprisingly, having a self-employed mother while growing up has no impact on the individual's self-employment choice. One conjecture for this finding is that many of these mothers may have been working in the family business with the father, suggesting the father self-employment variable is capturing the impact of growing up in an "entrepreneurial" family.

As discussed in Section 3, the estimated coefficients associated with the Big 5 personality variables in the reduced-form probit model presented in Table 2.4 reflect the relationship of these variables to both performance in self and paid-employment, as well as preferences for entrepreneurship. We now begin to untangle these 2 sets of effects. Our next step is to estimate the sectoral income regressions given by equations (2.3.5) and (2.3.6), including inverse mills-ratio terms generated from the reduced-form probit estimates. While the model can be identified through the assumption of joint normality of the error terms, we also incorporate exclusion restrictions to aid in identification. In particular, we assume that the father's and mother's self-employment status, as well as whether each was present in the home when the individual was a child, affects only the decision to be self-employed and not income as either an entrepreneur or a paid employee. We experimented with allowing father's and mother's self-employment status to enter the income regressions, but neither variable was statistically significant for paid-employees or the self-employed.

Estimates of the selection-corrected sectoral income regressions (2.3.5) and (2.3.6) are presented in Table 2.5. The coefficients on the inverse mills-ratios are small and statistically insignificant, suggesting that sample selection issues are not a problem for our estimates. The first column of the table shows that while individuals who are more open to new experiences are more likely to choose self-employment, this personality trait is associated with poorer performance in entrepreneurship, although the coefficient is not statistically significant. Among the other characteristics, extraversion has the strongest positive impact on entrepreneurial performance; the coefficient is statistically significant at the 10% level. This finding may not be surprising given that extroversion is associated with higher levels of sociability and assertiveness. For entrepreneurs who have a high level of interaction with customers, funders, and employees, such characteristics may be especially valuable in increasing sales, raising funding, or encouraging a high degree of employee effort. By contrast, the estimates in second column of the table suggest that extraversion does not have a similarly strong effect on performance as a wage worker. It may be that organizational structure in a larger firm reduces the scope for extraversion to have an impact.

Studies have found that individuals exhibiting a higher degree of conscientiousness are better performers, perhaps reflecting their organized and detail-oriented natures. We obtain similar results here: more conscientious wage workers earn significantly higher pay. The coefficient is even larger in magnitude in self-employment although it is not statistically significant. Finally, column (2) shows that individuals who are more trusting and altruistic earn significantly less in paid employment. In contrast, the agreeableness trait has virtually no impact on entrepreneurial performance, as measured by income.

With regard to other characteristics, comparison of the estimates in columns (1) and (2) show a much smaller educational gradient for entrepreneurial performance as compared to paid employment. For example, the estimates imply that a college graduate earns approximately 24% more than a high school dropout in self-employment (and the estimate is not significant), while the college premium is 55% for paid employees. Other studies have found similar results, and have argued the difference reflects the idea that education has much less signaling value in self-employment than in wage work since the entrepreneur knows his or her own productivity. It is also notable that the "marriage premium" is similar across sectors, perhaps reflecting the spouse's role in home production that allows the individual to spend more time in paid work. Finally, family background as measured by the father's educational attainment has a significant effect on entrepreneurial income, but the coefficient estimate is close to zero and insignificant for paid employees.

The final step of our analysis is to use the estimated income coefficients from Table 2.5 to construct the sectoral difference in expected income for each individual in the sample, $E[y_{iSE}] - E[y_{iPE}]$, and then include this quantity as an additional variable in the structural probit model defined by equation (2.3.4). Since we control for the impact of the personality variables on income by including $E[y_{iSE}] - E[y_{iPE}]$, the coefficients on the personality variables in the structural probit reflect preferences for entrepreneurship associated with the Big 5. The estimates of equation (2.3.4) presented in Table 2.6 show that individuals expecting to earn more in self-employment are more likely to choose to be entrepreneurs. Consequently, individuals with traits such as extraversion that have higher returns in self-employment than in paid employment will be more likely to become entrepreneurs because of this performance effect. However, there is no evidence that more extraverted individuals have a stronger preference for entrepreneurship. Turning to other components of the Big 5, we once again find that individuals who are more open to new experiences are more likely to be self-employed. In contrast to our reduced form findings in Table 2.4, by controlling for the expected sectoral income differential we are now able to interpret this finding as reflecting preference considerations associated with openness, rather than performance considerations. Finally, more agreeable or conscientious individuals appear to have less preference for self-employment, perhaps because of the riskier and more unstructured nature of entrepreneurship, but again these effects are not statistically significant. Overall, our findings suggest that individuals who are more open to new experiences have a stronger preference for entrepreneurship, which may reflect less risk aversion on their part. However, once self-employed openness does not have a positive impact on entrepreneurial performance. On the other hand, individuals who exhibit a greater degree of extraversion do not appear to have strong preference for self-employment, but if they do start their own business they are more likely to be successful.

A concern for our analysis is the potential endogeneity of the Big 5 personality measures. While we allowed for self-selection in the estimation of the sectoral income equations, it may be the case, for example, that individuals who enter self-employment become more open to new experiences, rather than vice versa. One might consider instrumenting for the personality characteristics in the self-employment choice and sectoral income regressions, it is not clear that five instruments are available in the MIDUS data. Consequently, our approach is to first determine whether endogeneity of the Big 5 personality should be of concern.

There is debate in the literature regarding the stability of the Big 5 over an individual's life-cycle. While one can imagine that these personality characteristics might still be forming in adolescence, the question for our findings is whether they still change in response to environmental factors, such as occupation or income, after age 25 (the youngest age in our sample). To investigate this possibility, we use the 1995 data from MIDUS I for our sample members to estimate regressions of the form:

$$P_{ik2004} = \varkappa_{0k} + \varkappa_{1k} SE_{i1995} + \varkappa_{2k} y_{i1995} + \varkappa_{3k} P_{ik1995} + Z_{i1995}\varkappa_{4k} + u_{ik2004} \quad (2.5.1)$$

where $P_{ikt}$ is the $k$'th personality trait measured in year $t$, $SE_{i1995}$ is an indicator for self-employment status in 1995, and $y_{i1995}$ is the 1995 income of individual $i$. The vector $Z_{i1995}$ includes the age and educational level of individual $i$ in 1995. The parameters $\varkappa_{1k}$ and $\varkappa_{2k}$ indicate the extent to which personality characteristics change in response to self-employment status or income.

Estimates of $\varkappa_{1k}$, $\varkappa_{2k}$, and $\varkappa_{3k}$ from equation (2.5.1) for each of the Big 5 personality characteristics are reported in Table 2.7. We cannot reject the hypothesis that $\varkappa_{1k} = 0$ in any of the regressions. In addition, for all of the five personality characteristics we find no evidence that the measured personality trait is affected by income. Overall, these estimates give us some confidence that the relationships that we estimate between the Big 5 and self-employment performance and preferences for entrepreneurship are not contaminated by endogeneity problems for the most part. Individuals do not appear to become more open to new experience in response to being self-employed, and higher income is not associated with increased extraversion.

## 2.6 Conclusion

Recent studies in the entrepreneurship literature have found a relationship between personality characteristics, as measured by the Big 5, and self-employment choice. This finding potentially reflects two factors: (a) the impact of personality on performance as an entrepreneur or paid employee; (b) preferences for entrepreneurship that are related to personality characteristics. For example, laboratory studies have found that individuals who score higher on the openness to new experiences dimension are less risk averse. This study attempts to distinguish between these performance and preference explanations by estimating a structural model of self-employment choice using data on 1100 adults from the MIDUS survey in the United States. Our findings confirm that individuals who are more open to new experiences are more likely to entrepreneurs. However, we find that this reflects preferences for being self-employed; there is actually a negative, though insignificant, relationship between openness and entrepreneurial performance. Conversely, we find that more extroverted individuals (who tend to be more assertive and sociable) tend to be significantly more successful entrepreneurs, as measured by income. However, extraversion does not appear to affect preferences for self-employment.

Our findings have potentially important implications for public policy toward entrepreneurship. Providing untargeted subsidies to encourage business formation may have the effect of encouraging potentially poor performing individuals to become entrepreneurs. Essentially, such policies may subsidize individuals' preferences for entrepreneurship. On the other hand, targeting subsidies at potentially high performers, who our findings suggest have higher levels of extraversion, may be quite successful in encouraging the entry of higher quality startups. Of course, our paper only examines one dimension of performance, self-employment income, and it would be useful to consider other metrics of performance, such as survival or job creation over time.

## 2.7 Tables

Table 2.1: Big Five Facets

| Characteristic | Facet |
|---|---|
| Openness to Experience | Fantasy, Aesthetics, Feelings, Actions, Ideas, Values |
| Conscientiousness | Competence, Order, Dutifulness, Achievement striving, Self-discipline, Deliberation |
| Extraversion | Warmth, Gregariousness, Assertiveness, Activity, Excitement seeking, Positive emotions |
| Agreeableness | Trust, Straightforwardness, Altruism, Compliance, Modesty, Tender-mindedness |
| Neuroticism (Emotional Stability) | Anxiety, Angry hostility, Depression, Self-consciousness, Impulsiveness, Vulnerability |

Table 2.2: Summary Statistics

| Variable | Mean | Std Deviation | observations |
|---|---|---|---|
| | | 2004 | |
| Age | 49.92 | 8.01 | 1110 |
| Self Employed | 0.22 | 0.41 | 1110 |
| % GED | 0.04 | 0.2 | 1110 |
| % high school | 0.2 | 0.4 | 1110 |
| % some college | 0.26 | 0.44 | 1110 |
| % graduate | 0.49 | 0.5 | 1110 |
| Income Last Year | 69990.35 | 44806.69 | 1110 |
| % married | 0.79 | 0.41 | 1110 |
| Number of Kids | 3.18 | 1.74 | 1110 |

Table 2.3: Self Employed vs Paid Employees

| Variable | Paid | SE | Difference | p value |
|---|---|---|---|---|
| | | 2004 Other Variables | | |
| Age | 49.22 | 52.49 | -3.26 | 0 |
| % GED | 0.04 | 0.04 | 0 | 0.74 |
| % high school | 0.21 | 0.2 | 0.01 | 0.73 |
| % some college | 0.27 | 0.24 | 0.03 | 0.42 |
| % graduate | 0.48 | 0.52 | -0.04 | 0.26 |
| Income Last Year | 68547.86 | 75247.28 | -6699.42 | 0.04 |
| % married | 0.78 | 0.82 | -0.05 | 0.12 |
| Number of Kids | 3.13 | 3.36 | -0.23 | 0.08 |
| | | 2004 Personality Variables | | |
| Agreeableness | 3.24 | 3.28 | -0.04 | 0.34 |
| Extraversion | 3.03 | 3.12 | -0.09 | 0.04 |
| Neuroticism | 2.05 | 1.99 | 0.06 | 0.21 |
| Conscientiousness | 3.45 | 3.47 | -0.01 | 0.7 |
| Openess to Experience | 2.95 | 3.04 | -0.09 | 0.01 |
| Observations | 871 | 239 | | |

Table 2.4: Reduced Form Probit Model of Self-Employment Choice

| | coefficient |
|---|---|
| Agreeableness 2004 | -0.086 |
| | (0.101) |
| Extraversion 2004 | 0.099 |
| | (0.100) |
| Neuroticism 2004 | 0.001 |
| | (0.074) |
| Conscientiousness 2004 | -0.058 |
| | (0.107) |
| Openess to experience 2004 | 0.230** |
| | (0.109) |
| High school | 0.098 |
| | (0.248) |
| Some College | 0.061 |
| | (0.246) |
| Graduate | 0.082 |
| | (0.244) |
| Age | 0.026 |
| | (0.065) |
| Age Square | 0.000 |
| | (0.001) |
| Married | 0.135 |
| | (0.112) |
| Years of Education Father | 0.023** |
| | (0.012) |
| Years of Education Mother | 0.003 |
| | (0.015) |
| Father SE | 0.379*** |
| | (0.103) |
| Mother SE | 0.014 |
| | (0.164) |
| Father present at age 14 | -0.425* |
| | (0.233) |
| Mother present at age 14 | 0.662 |
| | (0.565) |
| Constant | -3.518** |
| | (1.788) |
| Observations | 1107 |

Standard error in parenthesis

*p<0.1, **p<0.05, ***p<0.01

Table 2.5: Sectoral (log) Income Regressions

| Wage | Self Employees | Paid Employees |
|---|---|---|
| Agreeableness 2004 | -0.025 | -0.176*** |
| | (0.136) | (0.044) |
| Extraversion 2004 | 0.248* | 0.048 |
| | (0.138) | (0.044) |
| Neuroticism 2004 | 0.019 | 0.030 |
| | (0.096) | (0.033) |
| Conscientiousness 2004 | 0.211 | 0.099** |
| | (0.143) | (0.047) |
| Openess to experience 2004 | -0.184 | 0.070 |
| | (0.175) | (0.051) |
| High School | -0.069 | 0.205** |
| | (0.333) | (0.104) |
| Some College | 0.058 | 0.305*** |
| | (0.327) | (0.103) |
| Graduate | 0.236 | 0.553*** |
| | (0.320) | (0.103) |
| Age | -0.048 | 0.081*** |
| | (0.088) | (0.028) |
| Age Squared | 0.001 | -0.001*** |
| | (0.001) | (0.000) |
| Married | 0.245 | 0.229*** |
| | (0.159) | (0.049) |
| Years of Education Father | 0.039*** | 0.006 |
| | (0.014) | (0.005) |
| Years of Education Mother | 0.013 | 0.007 |
| | (0.020) | (0.006) |
| Constant | 10.217*** | 8.062*** |
| | (2.667) | (0.723) |
| Inverse Mills Ratio | -.101 | -.095 |
| | (0.405) | (0.247) |
| Observations | 239 | 871 |

Standard error in parenthesis

*p<0.1, **p<0.05, ***p<0.01

Table 2.6: Structural Probit Estimates of Self-Employment Choice

| | coefficient |
|---|---|
| Difference in Self Employment vs Paid Income | 0.498* |
| | (0.260) |
| Agreeableness 2004 | -0.158 |
| | (0.106) |
| Extraversion 2004 | -0.013 |
| | (0.108) |
| Neuroticism 2004 | 0.006 |
| | (0.073) |
| Conscientiousness 2004 | -0.109 |
| | (0.108) |
| Openess to experience 2004 | 0.330*** |
| | (0.123) |
| High School | 0.181 |
| | (0.243) |
| Some College | 0.096 |
| | (0.237) |
| Graduate | 0.171 |
| | (0.230) |
| Age | 0.077 |
| | (0.072) |
| Age Squared | -0.000 |
| | (0.001) |
| Constant | -3.658** |
| | (1.800) |
| Observations | 1110 |

Standard error in parenthesis

*p<0.1, **p<0.05, ***p<0.01

Table 2.7: Determinants of Big 5 Personality Characteristics in 2004

| Variable | Agreeableness 2004 | Extraversion 2004 | Neuroticism 2004 | Conscientiousness 2004 | Openess to Experience 2004 |
|---|---|---|---|---|---|
| Self Employment in 1995 | -0.024 | 0.025 | 0.007 | 0.002 | 0.042 |
| | (0.032) | (0.032) | (0.039) | (0.027) | (0.030) |
| Income in 1995 | -0.013 | -0.001 | -0.004 | 0.020 | -0.000 |
| | (0.015) | (0.015) | (0.018) | (0.013) | (0.014) |
| Age | 0.007*** | 0.007*** | -0.004** | 0.001 | 0.004*** |
| | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) |
| Agreeableness 1995 | 0.629*** | | | | |
| | (0.024) | | | | |
| Extraversion 1995 | | 0.715*** | | | |
| | | (0.022) | | | |
| Neuroticism 1995 | | | 0.572*** | | |
| | | | (0.023) | | |
| Conscientiousness 1995 | | | | 0.599*** | |
| | | | | (0.025) | |
| Openess to Experience 1995 | | | | | 0.704*** |
| | | | | | (0.025) |
| Observations | 1070 | 1070 | 1070 | 1070 | 1070 |

All regressions include age, education dummmies

Standard error in parenthesis

*p<0.1, **p<0.05, ***p<0.01

# Bibliography

**Almlund, Mathilde, Angela Lee Duckworth, James J. Heckman, and Tim Kautz.** 2011. "Personality Psychology and Economics." IZA Discussion Papers 5500, Institute for the Study of Labor (IZA).

**Anger, Silke, and Guido Heineck.** 2010. "The Returns to Cognitive Abilities and Personality Traits in Germany." *Labour Economics*, 17(3): 535 – 546.

**Antecol, Heather, and Deborah A. Cobb-Clark.** 2010. "Do Non-Cognitive Skills Help Explain the Occupational Segregation of Young People?" IZA Discussion Papers 5093, Institute for the Study of Labor (IZA).

**Barrick, Murray R., and Michael K. Mount.** 1991. "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis." *Personnel Psychology*, 44(1): 1–26.

**Barrick, Murray R., Michael K. Mount, and Judy P. Strauss.** 1993. "Conscientiousness and Performance of Sales Representatives: Test of the Mediating Effects of Goal Setting." *Journal of Applied Psychology*, 78(5): 715 – 722.

**Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro.** 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics*, 112(2): 537–579.

**Blanchflower, David G, and Andrew J Oswald.** 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics*, 16(1): 26–60.

**Boissiere, M., J. B. Knight, and R. H. Sabot.** 1985. "Earnings, Schooling, Ability, and Cognitive Skills." *The American Economic Review*, 75(5): 1016–1030.

**Borghans, Lex, Bart H H Golsteyn, James J Heckman, and Huub Meijers.** 2009. "Gender Differences in Risk Aversion and Ambiguity Aversion." *Journal of the European Economic Association*, 7(2-3): 649–658.

**Borghans, Lex, Bas ter Weel, and Bruce A. Weinberg.** 2008. "Interpersonal Styles and Labor Market Outcomes." *Journal of Human Resources*, 43(4): .

**Bouchard, T., and J. Propping.** 1993. *Twins as a Tool of Behavioral Genetics.*: John Wiley & Sons , Chichester, UK..

**Bowles, Samuel, Herbert Gintis, and Melissa Osborne.** 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature*, 39(4): 1137–1176.

**Carey, G.** 2003. *Human Genetics for the Social Sciences..*: Sage Publications,Thousand Oaks, CA..

**Cawley, John, James Heckman, and Edward Vytlacil.** 2001. "Three Observations on Wages and Measured Cognitive Ability." *Labour Economics*, 8(4): 419 – 442.

**Chamorro-Premuzic, Tomas, and Adrian Furnham.** 2003. "Personality Predicts Academic Performance: Evidence from Two Longitudinal University Samples." *Journal of Research in Personality*, 37(4): 319 – 338.

**Daly, Michael, Liam Delaney, and Colm P. Harmon.** 2009. "Psychological and Biological Foundations of Time Preference." *Journal of the European Economic Association*, 7(2-3): 659–669.

**Devine, Theresa J.** 1994. "Characteristics of Self-Employed Women in the United States." *Monthly Labor Review*, 117(3): 20–34.

**Dohmen, Thomas.** 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review*, 100(3): 1238–60.

**Evans, David S., and Boyan Jovanovic.** 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *The Journal of Political Economy*, 97(4): pp. 808–827.

**Evans, David S, and Linda S Leighton.** 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review*, 79(3): 519–35.

**Fairlie, Robert W., and Alicia M. Robb.** 2007. "Why Are Black-Owned Businesses Less Successful than White-Owned Businesses? The Role of Families, Inheritances, and Business Human Capital." *Journal of Labor Economics*, 25 289–323.

**Filer, Randall K.** 1986. "The Role of Personality and Tastes in Determining Occupational Structure." *Industrial and Labor Relations Review*, 39(3): pp. 412–424.

**Fortin, Nicole M.** 2008. "The Gender Wage Gap among Young Adults in the United States: The Importance of Money versus People." *Journal of Human Resources*, 43(4): .

**Goldberg, Lewis R.** 1993. "The Structure of Phenotypic Personality Traits." *American Psychologist*, 48(1): 26 – 34.

**Goldberg, L.R.** 1971. "A Historical Survey of Personality Scales and Inventories." In *Advances in Psychological Assessment*. ed. by Paul McReynolds: Palo Alto, CA: Science and Behavior Books, 293 – 336.

**Goldberg, L.R.** 1981. "Language and Individual Differences: The Search for Universals in Personality Lexicons." In *L. Wheeler (Ed.). Review of personality and social psychology: Vol.* 2.: Beverly Hills, CA: Sage., 141 – 165.

**Goldsmith, Arthur H, Jonathan R Veum, and William Alexander Darity.** 1997. "The Impact of Psychological and Human Capital on Wages." *Economic Inquiry*, 35(4): 815–29.

**Groves, Melissa Osborne.** 2005. "How Important is your Personality? Labor Market Returns to Personality for Women in the US and UK." *Journal of Economic Psychology*, 26(6): 827 – 841.

**Hamilton, Barton H.** 2000. "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment." *Journal of Political Economy*, 108(3): 604–631.

**Hampson, Sarah E., Lewis R. Goldberg, Thomas M. Vogt, and Joan P. Dubanoski.** 2006. "Forty Years On: Teachers' Assessments of Children's Personality Traits Predict Self-Reported Health Behaviors and Outcomes at Midlife." *Health Psychology*, 25(1): 57 – 64.

**Heckman, James J, and Paul A LaFontaine.** 2010. "The American High School Graduation Rate: Trends and Levels." *Review of Economics and Statistics*, 92(2): 244–262.

**Heckman, James J, and Yona Rubinstein.** 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review*, 91(2): 145–149.

**Heckman, James J., Jora Stixrud, and Sergio Urzua.** 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics*, 24(3): 411–482.

**Heineck, Guido.**"Does it Pay to be Nice? Personality and Earnings in the UK."

**Hettema, John, Michael Neale, and Kenneth Kendler.** 1995. "Physical Similarity and the Equal-Environment Assumption in Twin Studies of Psychiatric Disorders." *Behavior Genetics*, 25 327–335.

**Hipple, S.** 2010. "Self-Employment in the United States." *Monthly Labor Review*, 133(9): 17–32.

**Hogan, Joyce, and Brent Holland.** 2003. "Using Theory to Evaluate Personality and Job-Performance Relations: A Socioanalytic Perspective." *Journal of Applied Psychology*, 88(1): 100 – 112.

**Hogan, Robert, Joyce Hogan, and Brent W. Roberts.** 1996. "Personality

Measurement and Employment Decisions: Questions and Answers." *American Psychologist*, 51(5): 469 – 477.

**Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen.** 1994. "Sticking it Out: Entrepreneurial Survival and Liquidity Constraints." *The Journal of Political Economy*, 102(1): pp. 53–75.

**Hout, Michael, and Harvey Rosen.** 2000. "Self-Employment, Family Background, and Race." *The Journal of Human Resources*, 35(4): 670–692.

**Judge, Timothy A., Chad A. Higgins, Carl J. Thoresen, and Murray R. Barrick.** 1999. "The Big Five Personality Traits, General Mental Ability, and Career Success Across the Life Span." *Personnel Psychology*, 52(3): 621–652.

**Judge, Timothy A, and Charlice Hurst.** 2007. "Capitalizing on One s Advantages : Role of Core Self-Evaluations." *Journal of Applied Psychology*, 92(5): 1212–1227.

**Kanfer, Ruth, Connie R. Wanberg, and Tracy M. Kantrowitz.** 2001. "Job Search and Employment: A Personality Motivational Analysis and Meta-Analytic Review." *Journal of Applied Psychology*, 86(5): 837 – 855.

**Kendler, Kenneth S., Michael C. Neale, Ronald C. Kessler, Andrew C. Heath, and Lindon J. Eaves.** 1993. "A Test of the Equal-Environment Assumption in Twin Studies of Psychiatric Illness." *Behavior Genetics*, 23 21–27.

**Krueger, A, and D Schkade.** 2008. "Sorting in the Labor Market: Do Gregarious Workers Flock to Interactive Jobs." *Journal of Human Resources*, 92(8-9): 1833–1845.

**Lindqvist, Erik, and Roine Vestman.** 2011. "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment." *American Economic Journal: Applied Economics*, 3(1): 101–28.

**McCrae, Robert R., and Paul T. Costa.** 1987. "Validation of the Five-Factor Model of Personality Across Instruments and Observers." *Journal of Personality and Social Psychology*, 52(1): 81 – 90.

**McCrae, Robert R., and Paul T. Costa.** 1997. "Personality Trait Structure as a Human Universal." *American Psychologist*, 52(5): 509 – 516.

**McGue, Matt, Steven Bacon, and David T. Lykken.** 1993. "Personality Stability and Change in Early Adulthood: A Behavioral Genetic Analysis." *Developmental Psychology*, 29(1): 96 – 109.

**Moffitt, Terrie E., Avshalom Caspi, and Michael Rutter.** 2005. "Strategy for Investigating Interactions Between Measured Genes and Measured Environments." *Arch Gen Psychiatry*, 62(5): 473–481.

**Mueller, Gerrit, and Erik Plug.** 2006. "Estimating the Effect of Personality on Male and Female Earnings." *Industrial and Labor Relations Review*, 60(1): 3–22.

**Murnane, Richard J., John B. Willett, and Frank Levy.** 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *The Review of Economics and Statistics*, 77(2): 251–266.

**Nicolaou, Nicos, and Scott Shane.** 2009. "Can Genetic Factors Influence the Likelihood of Engaging in Entrepreneurial Activity?" *Journal of Business Venturing*, 24(1): 1–22.

**Nicolaou, Nicos, Scott Shane, Lynn Cherkas, Janice Hunkin, and Tim D. Spector.** 2008a. "Is the Tendency to Engage in Entrepreneurship Genetic?" *Management Science*, 54(1): 167–179.

**Nicolaou, Nicos, Scott Shane, Lynn Cherkas, and Tim D. Spector.** 2008b. "The Influence of Sensation Seeking in the Heritability of Entrepreneurship." *Strategic Entrepreneurship Journal*, 2(1): 7–21.

**Ones, Deniz S., Stephan Dilchert, Chockalingam Viswesvaran, and Timothy A. Judge.** 2007. "In Support of Personality Assessment in Organizational Settings." *Personnel Psychology*, 60(4): 995–1027.

**Paulson, Anna L., and Robert Townsend.** 2004. "Entrepreneurship and Financial Constraints in Thailand." *Journal of Corporate Finance*, 10(2): 229–262.

**Paulson, Anna L., Robert M. Townsend, and Alexander Karaivanov.** 2006. "Distinguishing Limited Liability from Moral Hazard in a Model of Entrepreneurship." *The Journal of Political Economy*, 114(1): 100–144.

**Plomin, Robert, J.C. DeFries, and John C. Loehlin.** 1977. "Genotype-Environment Interaction and Correlation in the Analysis of Human Behavior." *Psychological Bulletin*, 84(2): 309 – 322.

**Quadrini, Vincenzo.** 1999. "The Importance of Entrepreneurship for Wealth Concentration and Mobility." *Review of Income and Wealth*, 45(1): 1–19.

**Rauch, A, and M Frese.** 2007. "Let's Put the Person Back into Entrepreneurship Research: A Meta-Analysis on the Relationship between Business Owners' Personality Traits, Business Creation, and Success." *European Journal of Work and Organizational Psychology*, 16(4): 353–385.

**Robbins, Steven B., Jeff Allen, Alex Casillas, Christina Hamme Peterson, and Huy Le.** 2006. “Unraveling the Differential Effects of Motivational and Skills, Social, and Self-Management Measures from Traditional Predictors of College Outcomes.” *Journal of Educational Psychology*, 98(3): 598 – 616.

**Roberts, Brent W.** 2007. “Contextualizing Personality Psychology.” *Journal of Personality*, 75(6): 1071–1082.

**Roberts, Brent W., and Wendy F. DelVecchio.** 2000. “The Rank-Order Consistency of Personality Traits from Childhood to Old Age: A Quantitative Review of Longitudinal Studies.” *Psychological Bulletin*, 126(1): 3 – 25.

**Roy, A. D.** 1951. “Some Thoughts on the Distribution of Earnings.” *Oxford Economic Papers*, 3(2): 135–146.

**Sacerdote, Bruce.** 2007. “How Large Are the Effects from Changes in Family Environment? A Study of Korean American Adoptees.” *The Quarterly Journal of Economics*, 122(1): 119–157.

**Scarr, Sandra, and Louise Carter-Saltzman.** 1979. “Twin Method: Defense of a Critical Asumption.” *Behavior Genetics*, 9 527–542.

**Schmidt, Frank L., and John E. Hunter.** 1998. “The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings.” *Psychological Bulletin*, 124(2): 262 – 274.

**Sutin, Angelina R., Paul T. Costa, Richard Miech, and William W. Eaton.** 2009. “Personality and Career Success: Concurrent and Longitudinal Relations.” *European Journal of Personality*, 23(2): 71–84.

**Urza, Sergio.** 2008. "Racial Labor Market Gaps: The Role of Abilities and Schooling Choices." *Journal of Human Resources*, 43(4): .

**White, Roderick E., Stewart Thornhill, and Elizabeth Hampson.** 2006. "Entrepreneurs and Evolutionary Biology: The Relationship between Testosterone and New Venture Creation." *Organizational Behavior and Human Decision Processes*, 100(1): 21–34.

**White, Roderick E., Stewart Thornhill, and Elizabeth Hampson.** 2007. "A Biosocial Model of Entrepreneurship: The Combined Effects of Nurture and Nature." *Journal of Organizational Behavior*, 28(4): 451–466.

**Zhang, Zhen, Michael J. Zyphur, Jayanth Narayanan, Richard D. Arvey, Sankalp Chaturvedi, Bruce J. Avolio, Paul Lichtenstein, and Gerry Larsson.** 2009. "The Genetic Basis of Entrepreneurship: Effects of Gender and Personality." *Organizational Behavior and Human Decision Processes*, 110(2): 93 – 107.

**Zhao, Hao, and Scott E Seibert.** 2006. "The Big Five Personality Dimensions and Entrepreneurial Status: a Meta-Analytical Review." *Journal of Applied Psychology*, 91(2): 259–271.

**Zhao, Hao, Scott E. Seibert, and G. Thomas Lumpkin.** 2009. "The Relationship of Personality to Entrepreneurial Intentions and Performance: A Meta-Analytic Review." *Journal of Management*.

Chapter 3

# Using a Model of Location Around the Clock to Recover the Amenity Value of Daylight

## 3.1 Introduction

We investigate how the timing of activities is determined. Rather than focusing on how much time people devote to some activities, we try to understand at what specific hours those activities are undertaken. In other words, assuming that individuals sleep before work and have leisure after it, we basically study work shift choice. Prominent sociologists such as Presser (2004) have pointed out that wage premia are very small for night shift workers and thus cannot be a significant determinant of night shift choice. However, the observed wage premia is a difference between averages computed over highly selected samples. This simple average ignores that night shift workers usually come from the bottom of the conditional wage distribution (i.e. they have some unobservable characteristics that make them have lower earnings than otherwise observationally identical workers). In any event, night shift appears to be a disamenity and it is therefore chosen by people with fewer day shift opportunities. Building upon the framework outlined by Rosen (1986) which emphasizes the role of amenities and disamenities on wages, Kostiuk (1990) recognizes the importance of this selection issue and finds more significant night shift premia. His analysis resembles that of the classic textbook mover/stayer model of migration. In fact, one can think of night shift work as a migration decision where the individual migrates in a temporal rather than a spatial dimension. Then, it should not come as a surprise that similar tools can be applied for its analysis. The focus here differs from Kostiuk's in that we try to emphasize the role of social interactions in explaining shift choice. In particular, we claim that a distribution of workers (and jobs) highly concentrated on conventional office hours may arise as a suboptimal or inefficient equilibrium in an economy with multiple equilibria.

The paper asks the following related questions: When do we work? What's the amenity value of day light? Do most people sleep at night because it is dark or because most other people do that? Assuming that most people demand a compensating wage differential to supply their labor services at night, how this premium changes in alternative scenarios in which the share of the population that works at night is substantially higher? If darkness is not so important should we all coordinate and work at the same time?

We conjecture that the economy displays multiple equilibria. To the extent that lack of daylight is not the only determinant of shift premia, the current equilibrium might be inefficient.

The remainder of this paper is organized as follows: Section 3.2 briefly overviews related literature in economics and sociology on "non-standard" work hours. Section 3.3 describes the data and presents descriptive evidence on the prevalence of shift work as well as estimates of the shift premium that corrects for selection into shift. Section 3.4 follows the work of Bayer and Timmins (2005, 2007) in developing a structural equilibrium sorting model of location around the 24 hour clock that allows for social interactions in a consistent way. Section 3.5 describes the estimation strategy and Section 3.6 presents the results and explores willingness to pay for daylight. Conclusions follow.

## 3.2 Related Literature

Our paper is related to several strands of literature. First, there is the issue of whether there is a shift premium at all. Modern empirical economists are firmly rooted in the Rosen (1986) compensating differential framework and so they tend to see shift work as another job disamenity which should command a premium in a hedonic equilibrium.[1] More sociological perspectives tend to question the existence of such shift premia.[2] However, work by Kostiuk (1990) has showed that comparing the same occupations and controlling for selection into shift is very important to identify shift premia.[3]

Second, since we conjecture that one of the reasons people dislike working at night has to do with loneliness or the inability to enjoy leisure time jointly with others, our paper is related to a literature on leisure externalities. In particular, Alesina, Glaeser and Sacerdote (2005) who emphasize the social multiplier effects of European labor market policies that tend to reduce work hours and increase the employment rate. Jenkins and Osberg (2005) use British data to document the

[1] See also Khan (2008).

[2] See Presser(2004) for a comprehensive analysis of non-standard work hours from a sociological perspective.

[3] See also Hwang, Reed and Hubbard (1992) and Hamermesh (1999) for the implications of unobservables when estimating shift premia.

important externalities associated with working time decisions. They find that an individual's time use choices are contingent on the time use choices of others because the utility derived from leisure time often benefits from the presence of companions inside and outside the household.

Given our emphasis on externalities, our paper is also related to a broad literature on the econometrics of social interactions. This literature surfaces with different names in different fields of economics: peer effects in education, agglomeration and congestion effects in urban economics or network externalities in industrial organization and marketing. Our work is more closely related to Brock and Durlauf (2001) and in particular, the urban approach of Bayer and Timmins (2005, 2007). In our model, the utility from choosing a particular location around the clock will depend on the share of individuals choosing that same clock location.

Finally, an extensive literature on firm task scheduling take positive shift premia for evening and night shift as given and outlines their implications for the organization of production and tasks around the clock within firms.[4]This literature emphasizes the productive loss implicit in idle capital and the associated distortion on firm size. Coupled with results from this literature, our estimates could be used to derive potential welfare gains from moving to a more evenly distributed work force around the clock. Moreover, the more efficient use of resources would not limited to the capital stock within productive firms but would also apply to public infrastructure more generally.[5]

[4]See Marris (1964), Georgescu-Roegen (1970), Winston (1974), Betancourt and Clague (1981), Betancourt (1986). See also Calvo (1975) for a more macroeconomic perspective on capital idleness and Weiss (1996) for a therethical analysis of synchronization of work schedules.

[5]Urban planners have envisioned the potential welfare gains from this type of around-the-clock scenario. See the urban planning literature on "compact city" spurred by the seminal work of Dantzig and Saaty (1973).

## 3.3 The Data

### 3.3.1 Decennial Population Census

We rely heavily on publicly available microdata from the Decennial Population Census of 2000. In particular, we exploit some of the information provided by the Journey to Work module in the long form questionnaire. The following two questions allow us to compute measure of shift choice and more generally, measures of location around the clock.

- At what time you leave to work ?
- How long it takes you to get to work ?

We also exploit standard demographic covariate information available in the Census including race, gender, marital status, etc. Measures of labor earnings and hours of work allow us to identify full time workers and construct estimates of hourly wage. Occupational classification codes allow us to compare wages for workers in the same type of jobs. In addition, our empirical strategy below exploits geographic variation in exposure to daylight conditional on shift choice. In fact, the average daily exposure to daylight differs across space for a given choice of location around the clock. We rely on PUMAs (Public Use Microdata Areas) and the associated latitude and longitudes of their geographic centroids to measure the spatial location of each individual. Sometimes, several PUMAs are close enough to each other that we construct bundles of PUMAs (essentially combinations of

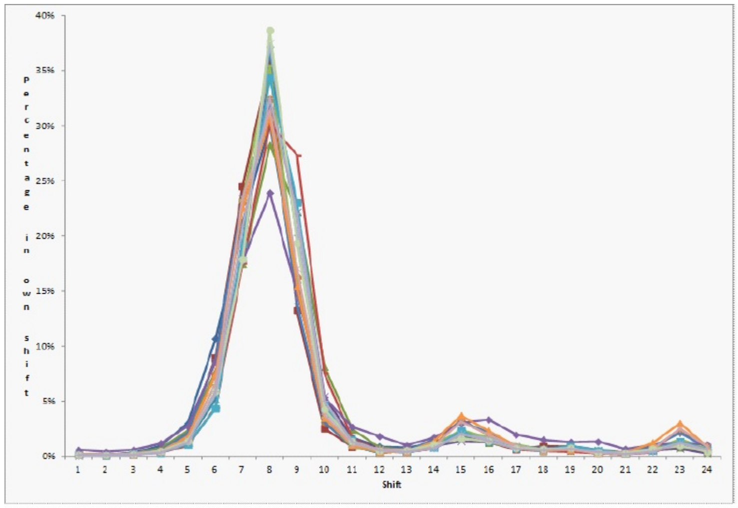


Figure 3.1: Distribution of Workers Across Shifts in each Bundle

several PUMAs adjacent to each other) to base our analysis of social interactions on a common notion of geographic distance.[6]

We keep a small number of observations for our final analysis sample. The Appendix provides details on the construction of the sample. Table 3.1 presents descriptive statistics for our main variables. Key to our social interaction analysis, we also observe what the rest of the population is doing in the same geographic location. We define spatial bundles of PUMAs in a given radius around each sampled person. We describe the construction of bundles of PUMAs in the Appendix. Figure 3.1 shows the distribution of location choices around the 24-hour clock for the bundles in our analysis sample. In particular, the figure shows the distribution of times at which individuals begin full time work.

[6]Since PUMAs are delimited by the number of sampled individuals, PUMAs for densely populated areas are relatively small. Therefore, we often times construct bundles of PUMAs to preserve a homogenous notion of distance relevant for social interactions.

As can be seen in the figure, while the population is heavily concentrated in and around location 8, some non-negligible fraction begins work at some other non-standard hours. In particular, the profile display some bumps at clock locations 15 and 23 mostly associated with the beginning of the second and third shift. Moreover, note that the distribution of workers around the clock exhibits some heterogeneity across bundles.

### 3.3.2 Sunset and Sunrise Times

In order to measure exposure to daylight associated with different shift choices in each bundle we rely on bundle-specific sunrise and sunset times. We leverage an astronomical formula that allow us to determine the sunrise and sunset time, and resulting hours of daylight available for each day of the year in a given latitude and longitude. We then apply this to the (lat,long) pair for each location (i.e. at the centroid of each bundle of PUMAs). Using this information we then derive exposure to daylight associated with each shift choice by assuming that workers sleep for eight uninterrupted hours, wake up about two hours before beginning to work and remain awake for sixteen hours before falling asleep again.[7] We are then able to compute, for each possible clock location in each spatial bundle, how many of those sixteen hours involve exposure to daylight. Finally we average, these shift-specific daily exposures to daylight across the year. Note that there is substantial latitudinal variation in availability of daylight across the year. For example the amount of daylight in South Florida is fairly constant across the year, whereas in places like

[7] We abstract from the choice of number of hours of sleep. There is a surprisingly small literature in economics that looks at the important issue of optimal choice of hours of sleep. See Bergstrom (1977), Hoffman (1977), Biddle and Hamermesh (1990), Hamermesh (2002) and Yaniv (2004)

Seattle or Minneapolis is large during the Summer months and small during the Winter. Moreover, conditional on a given latitude, there is substantial longitudinal variation generated by time zones.[8] For example, consider two individuals who wake up at 6am according to their own time zones and who are living in two cities with common latitude. Say, one of the cities is located at the west end of the eastern time zone whereas the other is located at the east end of the central time zone. Even when they are relatively close to each other and they wake up at the "same" time, their exposure to daylight will be different because the individual in the Central time zone is effectively waking up one hour earlier and this induce a different level of exposure.

Figure 3.2 documents the variation in exposure to daylight across different location around the clock for our spatial bundles. Of course there is large variation across shifts within bundles. For example, somebody who begins to work at 9am is assumed to wake up at 7am and go to sleep at 11pm. This implies an exposure of 11 to 12 hours of daylight on average across the year. On the other hand, somebody starting work at 11pm is in general exposed to approximately 6 hours of daylight. Note that there is also some variation across bundles for a given shift.

### 3.3.3 Empirical Evidence on Shift Premia

Before turning to the model we examine whether in a first look at the data we observe any evidence of shift premia. We create and indicator of whether an individual is a day worker $(d = D)$ or a shift worker $(d = S)$. We use information on the time

[8] See Hamermesh et al. (2008) for an explicit analysis of the role of time zones and television schedules on patterns of work and sleep timing.

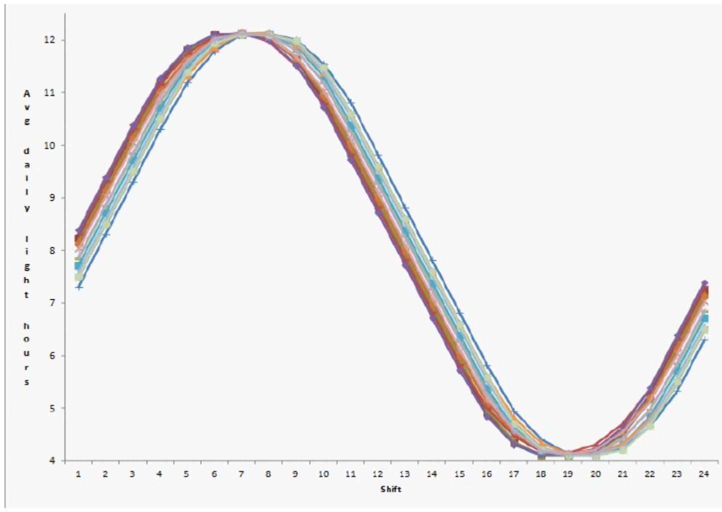


Figure 3.2: Daily Exposure to Daylight Across Shifts and Bundles of PUMA's

the individual departs from home to work and the time it takes him to get to work to come up with an estimated time at which he starts working. Given the time at which he starts working and the number of hours he works, we can compute the number of hours the the individual works in each of the following two time ranges.

1. 8am to 4pm

2. 4pm to 8am.

We then define $d_i = D$ (he/she is a day worker) if most of his work hours are in the 8am-4pm range and we define $d_i = S$ (he/she is a shift worker) if most of his hours fall in the 4pm-8am range.

The wage equation for individual $i$ in bundle $b$ when he chooses to work during the day is given by

$$\log(W_{iDb}) = X_i\beta_D + \lambda_{Db} + \varepsilon_{iDb} \tag{3.3.1}$$

whereas if he chooses to be a shift worker it is given by

$$\log(W_{iSb}) = X_i\beta_S + \lambda_{Sb} + \varepsilon_{iSb} \tag{3.3.2}$$

$X_i$ includes experience, education, gender, race, marital status, veteran status, an indicator of whether the individual lives in a statistical metropolitan area (SMSA) and indicators for different occupations. $\lambda_{Db}$ and $\lambda_{Sb}$ denote indicators for the different spatial bundles. Of course, we only get to see each individual working either as a day worker or shift worker and therefore can only hope to observe his wages in only one of these statuses. So we consider selection into shift work and day work by estimating a selection model

$$d_{ib} = \begin{cases} D & \text{if } Z_i\pi + \eta_{ib} > 0 \\ S & \text{otherwise} \end{cases} \tag{3.3.3}$$

where $Z_i = (X_i, Z_{1i})$ and $Z_{1i}$ serves as an exclusion restriction that explains selection into shift but is excluded from teh wage equation. In the empirical implementation $Z_{1i}$ denotes the percentage of shift workers within the industry in which individual $i$ works. We use this to selection correct the log wage equations and obtain estimates $(\widehat{\beta}_D, \widehat{\beta}_S)$. Table 3.2 presents the results

The coefficient associated with the inverse mills ratio is highly significant, implying strong selection of workers into day and shift work. Using these estimates we can compute predicted wages for day and shift work $(\widehat{W}_{iD}, \widehat{W}_{iS})$ for every individual and obtain estimates of the individual-specific shift differential as $100 \times \frac{\widehat{W}_{iS} - \widehat{W}_{iD}}{\widehat{W}_{iD}}$.

Then we can compute the average shift differential across the whole sample of workers or for particular subsamples by simply averaging these individual specific shift differentials.

Table 3.3 presents the results. Column 1 shows simple OLS estimates. Column 2 shows our preferred estimates that correct for selection into shift. As can be seen, the shift premia is much smaller when we fail to control for selection into shift work. The shift premium ranges between 3 to 5 percent. Once we account for selection we find a shift premium of approximately 39%.

As can be seen, a large shift premia is necessary to elicit labor supply at non-standard hours and in equilibrium the bulk of the labor force continues to offer its services at conventional hours. But really, is human productivity that much higher during the day? Most likely not. Why so few people work on night shifts? It seems clear that people prefer day shift. But, is this mostly because they wish to avoid the exposure to darkness associated with night shift or is it because almost everybody else is working during the day? If the latter, we may in fact be stuck in a bad (inefficient) equilibrium: everybody is behaving optimally, conditional on what others are doing, but as a whole the equilibrium is not efficient. This a relatively complex empirical question. In the next section we present the model that we will estimate to attempt an answer.

## 3.4 A Model of Shift Choice

We adapt the urban framework in Bayer and Timmins (2005, 2007) to our context by formulating a model of sorting around the clock. A location in our context is a specific hour around the clock in which the individual decides to begin to work every day. Unlike the binary shift formulation used by Kostiuk (1990), the hourly specification with 24 choices provides more precision when it comes to measure exposure to light.

We take spatial location $b$ as exogenous and model the choice of location around the clock $j$. A location around the clock is indexed by the time an individual starts work and it entails an uninterrupted period of 16 hours that starts when he wakes up (two hours before starting to work) and ends when the individual goes to bed.

Consider individuals living in spatial bundles $b = 1, ...., B$.[9].

$$U_{ijb} = X'_{jb}\beta_i + \alpha_i\sigma_{jb} + \gamma W_{ijb} + \xi_{jb} + \varepsilon_{ijb} \tag{3.4.1}$$

where $X_{jb}$ is a vector of exogenous attributes of clock location $j$ in bundle $b$. A key attribute we are focusing on is daily exposure to daylight (in hours on average across the year). $\sigma_{jb}$ is the fraction of workers who choose clock location $j$ in spatial bundle $b$. This is the measure that will help us capture social interaction effects, if any. $\xi_{jb}$ is an attribute for clock location $j$ in bundle $b$ that is observed by individuals making clock location decisions but unobserved by us, the econometricians. It collects and summarizes what individuals observe about each clock location and affects their utility. $\varepsilon_{ijb}$ represent i.i.d. (across individuals and clock locations) taste shifters with extreme value distribution.

$\alpha_i$ captures the strength and sign of the social interaction for each individual $i$.

[9] Explain how we define these bundles of MSAs. Centroid. Radius. etc...to get at the geographic reference point for each individual.

When $\alpha_i < 0$ there is a congestion effect. In this case, individuals tend to dislike very popular clock locations. When $\alpha_i > 0$ there is an agglomeration effect. Individuals tend to enjoy popular locations per se, over and above the features (if any) that make them popular. We allow the social interaction effect to vary with exogenous observable individual demographics $Z_i$ according to

$$\alpha_i = \alpha_0 + Z_i'\alpha_1 \tag{3.4.2}$$

$\beta_i$ is an individual specific taste parameter that can also vary with exogenous observables $Z_i$. It captures the marginal utility of additional hours of exposure to daylight.

$$\beta_i = \beta_0 + Z_i'\beta_1 \tag{3.4.3}$$

and $W_{ijb}$ is the Real Hourly Wage for individual $i$, if choosing clock location $j$ in spatial bundle $b$.

Replacing (5) and (6) into (4) we obtain

$$\begin{aligned} U_{ijb} &= X_{jb}'(\beta_0 + Z_i'\beta_1) + (\alpha_0 + Z_i'\alpha_1)\sigma_{jb} + \gamma W_{ijb} + \xi_{jb} + \varepsilon_{ijb} \\ U_{ijb} &= \delta_{jb} + X_{jb}'Z_i\beta_1 + \sigma_{jb}Z_i\alpha_1 + W_{ij}\gamma + \varepsilon_{ijb} \end{aligned}$$

$$\delta_{jb} = X_{jb}'\beta_0 + \alpha_0\sigma_{jb} + \xi_{jb}$$

We consider the following assumptions: $U_{ijb}$ does not depend on $X_{kb}$ or $\sigma_{kb}$ for $k \neq j, \sigma_{jb}$ is the only endogenous variable, $\overline{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, ....., \varepsilon_{iJ})$ is observed by all $i = 1, ....., N$ and individuals play a static simultaneous move game and equilibrium behavior is governed by a simple Nash Equilibrium concept. A continuum of individuals exist for each observed realization of $Z_i$. This allows us to integrate out unobserved preferences and work directly with choice probabilities.

We can then write the clock location choice probabilities as

$$P_{ij} = g_{ij}(Z_i, \overline{X}, \overline{\sigma}; \theta) \ \forall i, j$$

where

$$\begin{aligned}
\sigma_j &= \sum_i P_{ij} = \sum_i g_{ij}(Z_i, \overline{X}, \overline{\sigma}; \theta) \\
\sigma_j &= g_j(\overline{Z}, \overline{X}, \overline{\sigma}; \theta) \quad \forall j
\end{aligned}$$

$$\overline{\sigma} = g(\overline{Z}, \overline{X}, \overline{\sigma}; \theta)$$

$$\theta = \{\alpha, \beta, \gamma, \xi\}$$

**Definition 1** (Sorting Equilibrium)**.** *A sorting equilibrium is a set of individual shift work decisions that are each optimal given the shift work decisions of all other individuals in the population.*

Note that a fixed point for $\overline{\sigma} = g(\overline{Z}, \overline{X}, \overline{\sigma}; \theta)$ satisfies the above definition and it is easy to show that by Brower's Fixed Point theorem, a fixed point to the mapping defined by $g$ exists. Equilibrium may or may not be unique, though. Whether or not the equilibrium is unique depends on

1. the sign and magnitude of the social interaction, $\alpha$

2. the meaningful variation in household tastes $\sigma^2_\beta$;

3. the meaningful variation in fixed attributes across shift choices; and

4. the total number of choices (in this case $J = 24$)

The economic intuition for uniqueness goes as follows. Congestion effects preserve the collective rank order of locations. Same applies to mild $(0 < \alpha < \overline{\alpha})$ agglomeration effects. On the other hand, strong agglomeration effects have the ability to alter the equilibrium rank order of locations and thus allow for multiple equilibria. Since no closed form for $\overline{\alpha}$ is available, Bayer and Timmins (2005) show how a simulation based approach can be used to computationally explore how changes in the characteristics of the shift choice problem would affect the likelihood of multiple equilibria (how they affect the threshold $\overline{\alpha}$). They conclude that the larger the number of choices $J$, the meaningful variation in exogenous shift characteristics, $\beta_0\sigma^2_X$ or the heterogeneity in household preferences, $\sigma^2_\beta$ , the larger the maximum agglomeration effect that can sustain uniqueness. The basic economic intuition for these results is that increases in any of the three dimensions reduce the ability of a given agglomeration effect to alter the rank-order of collective preferences. Importantly, the estimator proposed by Bayer and Timmins (2007) is robust to the presence of multiple equilibria

## 3.5 Estimation

We follow Bayer and Timmins (2007) in using a two step strategy. In the first step, we proceed by maximum likelihood on the clock location choice microdata. In the second step, 2SLS is used to decompose the shift specific effects, $\{\delta_{jb}\}_{j=1}^{J}$ for each bundle $b$.

The empirical strategy exploits spatial variation in sunset and sunrise timing data combined with Population Census microdata on earnings and shift choice data.[10]

Recall clock location preferences are given by

$$U_{ijb} = \delta_{jb} + X'_{jb} Z_i \beta_1 + \sigma_{jb} Z_i \alpha_1 + W_{ij}\gamma + \varepsilon_{ijb} \tag{3.5.1}$$

$$\delta_{jb} = X'_{jb}\beta_0 + \alpha_0 \sigma_{jb} + \xi_{jb} \tag{3.5.2}$$

Using sunset and sunrise timing data we compute $\{\{X_{jb}\}_{j=1}^{J}\}_{b=1}^{B}$ the average exposure to daylight associated with choosing clock location $j$, when living in spatial bundle $b$. We exploit the Census data on earnings to compute the hourly wage that each individual would earn in the event of choosing shift $j$, when living in bundle $b$. Of course, we only observe $W_{ij}$ for individual $i$ and his chosen clock location $j$ but not at other clock locations. We need to predict what those counterfactual wages would be. For this we estimate the following wage equations for each clock location $j = 1, 2....J$ correcting for selection into it.

$$\log(W_{ijb}) = \theta_0 + \theta_1 Z_i^W + \theta_j D_{ij} + \lambda_b + \eta_{ij} \tag{3.5.3}$$

where $D_{ij}$ is an indicator that equals one when the individual chooses clock loca-

[10]Not only do we use the shift choice microdata in its own right but we also aggregate the microdata on shift choice to come up with shares of population choosing to "wake up" at different times in each spatial location.

tion $j$ and it is zero otherwise. $Z_i^W$ are individual determinants of wages such as education, experience, gender, race, veteran status. We also control for occupation effects given that some low wage occupations are disproportionately concentrated in night shifts. In addition we allow for bundle dummies, $\lambda_b$. As exclusion restriction, in the selection equation we include an indicator of whether the individual is single and has at least one child. Using these estimates, for each individual $i$ in bundle $b$ we predict $\widehat{W}_{ijb}$ for all $j$

After estimating these unobserved we proceed with the two-step strategy. First, we maximize the likelihood of the shift choice microdata. In doing so we estimate the first stage parameters

$$\theta_1 = \left( \left\{ \{\delta_{jb}\}_{j=1}^{J} \right\}_{b=1}^{B}, \alpha_1, \beta_1, \gamma \right) \tag{3.5.4}$$

where $\{\delta_{jb}\}_{j=1}^{J}$ is treated as a set of fixed shift effects for spatial bundle $b$. These fixed effects are estimated jointly with the other model parameters. While the model may display multiple equilibria, we can condition on the realized equilibrium and consistently recover the parameters from the microdata.

Indeed, if we let $\Omega = \{X, Z, W, [\delta, \alpha_1, \beta_1, \gamma]\}$ then the likelihood function is actually given by

$$L = \sum_{\sigma} P(\sigma|\Omega) \left\{ \prod_i \prod_j P_i(j|\sigma, \Omega)^{I_{i,j}} \right\} \tag{3.5.5}$$

and the probability that a given equilibrium arises, $P(\sigma|\Omega)$, does not depend on the particular location decisions of individual agents. We then avoid the complex part of the likelihood $P(\sigma|\Omega)$ and estimate parameters based on

$$\left\{\widehat{\delta}, \widehat{\alpha}_1, \widehat{\beta}_1, \widehat{\gamma}\right\} = \arg \max_{\delta,\alpha_1,\beta_1,\gamma} \left\{\prod_i \prod_j \Pr\left(d_{ij} = 1|\sigma, \Omega, Z_i\right)^{d_{ij}}\right\} \tag{3.5.6}$$

$$\Pr\left(d_{ij} = 1|\sigma, \Omega, Z_i\right) = \frac{\exp\left(\delta_{jb} + X'_{jb}Z_i\beta_1 + \sigma_{jb}Z_i\alpha_1 + \widehat{W}_{ijb}\gamma\right)}{\sum \exp\left(\delta_{kb} + X'_{kb}Z_i\beta_1 + \sigma_{kb}Z_i\alpha_1 + \widehat{W}_{ikb}\gamma\right)} \tag{3.5.7}$$

A bundle specific contraction mapping similar to that developed by Berry, Levinsohn and Pakes (1995) and adapted to urban choice models in Bayer, McMillan & Rueben (2011) is used to get $\{\delta_{jb}\}$ for each trial of $(\alpha_1, \beta_1, \gamma)$. Formally, the estimation problems becomes

$$\left\{\widehat{\delta}, \widehat{\alpha}_1, \widehat{\beta}_1, \widehat{\gamma}\right\} = \arg \max_{\alpha_1,\beta_1,\gamma} \left\{\prod_i \prod_j \Pr\left(d_{ij} = 1|\sigma, \Omega, Z_i\right)^{d_{i,j}}\right\} \tag{3.5.8}$$

$$\begin{aligned}
\text{s.t. } \{\delta_{jb}\}_{j=1}^{J} &= \text{the fixed point of the following contraction mapping} \\
&\quad \text{for bundle} b \\
\delta_{jb}^{(n+1)} &= T\left(\delta_b^{(n)}\right) \text{ for all } j, \text{ in bundle } b \\
T\left(\delta_b^{(n)}\right) &= \delta_{jb}^{(n)} - \ln\left(\frac{\widehat{\Pr}\left(d_j = 1|\alpha_1, \beta_1, \gamma, \left\{\delta_{kb}^{(n)}\right\}_{k=1}^{J}\right)}{\sigma_{jb}^{obs}}\right)
\end{aligned}$$

After recovering the first stage parameters 2SLS is used to decompose the shift specific effects, $\widehat{\delta}_{jb}$ by pooling data from all spatial bundles and using $\widehat{\delta}_{jb}$ as dependent variables in the following linear regression model

$$\widehat{\delta}_{jb} = X'_{jb}\beta_0 + \alpha_0\sigma_{jb} + \lambda_b + \xi_{jb} \tag{3.5.9}$$

OLS estimation of $\alpha_0$ would be inconsistent because of the endogeneity $\sigma_{jb}$. Indeed, the model implies that $\sigma_{jb}$ and $\xi_{jb}$ will most likely be correlated: each individual observes $\xi_{jb}$ and influences his clock location choice. Other individuals do the same. In the aggregate the share of individuals in spatial bundle $b$ choosing clock location $j$ will depend on $\xi_{jb}$. We then estimate $\alpha_0$ using IV to account for endogeneity of $\sigma_{jb}$.

Our proposed instrument exploits the fixed exogenous attributes of other clock location choices $k \neq j$. Indeed, this is plausibly valid instrument because clock location choice depends not only on own clock location attributes but also on how those attributes compare against the attributes of other clock locations. However, attributes of other clock locations $X'_{kb}$ for $k \neq j$ do not influence $\delta_{jb}$. We follow Bayer and Timmins (2007) and use the hypothetical share for each clock location that would arise absent any spillover if only observed attributes mattered.

To compute our instrument we consider the predicted clock location share $\widehat{\sigma}_{jb}$ that can be obtained from the model with estimated first stage parameters $\left\{\left\{\widehat{\delta}_{jb}\right\}_{j=1}^{J}\right\}_{b=1}^{B}, \widehat{\gamma}, \widehat{\beta}_1$ and initial guess $\widetilde{\beta}_0$ for $\beta_0$ but setting both the unobserved attribute and the social interaction effect to zero.

$$\xi = 0, \alpha = 0$$

The share given by the model is

$$\begin{aligned}
\sigma_{jb} &= \Pr\left(d_j = 1 | X, \alpha = 0, \xi = 0, \beta_0, \beta_1, \gamma\right) \\
&= \int \Pr\left(d_j = 1 | X, \alpha = 0, \xi = 0, \beta_0, \beta_1, \gamma, Z_i\right) f\left(Z_i\right) dZ_i
\end{aligned}$$

and can be estimated by

$$\begin{aligned}
\widehat{\sigma}_{jb} &= \widehat{\Pr}\left(d_j = 1 | X, \alpha = 0, \xi = 0, \widetilde{\beta}_0, \widehat{\beta}_1, \widehat{\gamma}\right) \\
&= \frac{1}{N_b} \sum_{i=1}^{N_b} \widehat{\Pr}\left(d_j = 1 | X, \alpha = 0, \xi = 0, \widetilde{\beta}_0, \widehat{\beta}_1, \widehat{\gamma}, Z_i\right) \\
&= \frac{1}{N_b} \sum_{i=1}^{N_b} \left[ \frac{\exp\left(X'_{jb}\left(\widetilde{\beta}_0 + Z'_i \widehat{\beta}_1\right) + \widehat{\gamma} W_{ij}\right)}{\sum_{k=1}^{J} \exp\left(X'_{kb}\left(\widetilde{\beta}_0 + Z'_i \widehat{\beta}_1\right) + \widehat{\gamma} W_{ik}\right)} \right]
\end{aligned}$$

We then obtain $\widehat{\beta}_0, \widehat{\alpha}_0$ by estimating by 2SLS given the guess $\widetilde{\beta}_0$ and using $\widehat{\sigma}_{jb}$ as instrument for $\sigma_{jb}$. We repeat this procedure until the guess $\widetilde{\beta}_0$ equals $\widehat{\beta}_0$. Standard errors are computed using bootstrap.[11]

### 3.6 Willingness to Pay for an Extra Hour of Daylight

In this section we use the model to derive measures of willingness to pay. These measures, in turn help us decompose the observed wage premium for shift work into a portion attributable to social interactions and a portion due to reduced exposure

[11]In the empirical implementation we actually focus on a subsample of individuals who work on occupations that have enough incidence across all the 24 locations around the clock so that the model can be identified. Therefore we split $\sigma_{jb}$ into a part that is endogenous and reflect aggregate clock location choices of individuals in spatial bundle $b$ and a completementary part that is exogenous and composed of workers who do not choose clock location. Therefore only the endogenous portion of $\sigma_{jb}$ is instrumented in our estimation algorithm.

to daylight. Table 3.4 presents the estimated parameters for a model specification in which we allow $\alpha_i$ and $\beta_i$ to vary by gender by considering $Z_i = 1$ if $i$ is a female, 0 otherwise.

With the estimates of the model at hand we can explore what's the willingness to pay for daylight in the population of workers. We can derive measures of willingness to pay for daylight based on $\widehat{\beta}_i$ and $\widehat{\gamma}$. Table 3.5 shows the results.

The results show that males have higher willingness to pay for daylight. As we can see the OLS estimates that ignore the endogeneity of the fraction of people in each clock location underestimate the willingness to pay for daylight by up to 20 percent. The preferred IV estimates show that males are willing to give up to 45 cents of their hourly wage for each additional hour of daily exposure to daylight. This means they are willing to pay up to \$3.59 per day (i.e. approximately \$18 per week or \$900 per year) for each additional hour of daylight. This implies, for example, that the daylight component of the shift premium required to induce somebody working in the 9am shift to work at the 11pm shift (for a reduction of $12 - 6 = 6$ hours of daylight exposure) would be approximately 18%[12]

We can also compute similar estimates of willingness to pay for "companionship". That we can estimate how much individuals are willing to pay to have an additional 1% of the population within a 10 mile radius to choosing their same clock location. Table 3.6 presents the results.

The estimates of WTP for companionship are large in magnitude. Here OLS overestimates the magnitude. Our preferred IV estimates indicate that males are

[12]Taking the average hourly wage in the sample \$15.08 we get ($\frac{0.45\times6}{14.6} = 0.18$)

willing to pay \$780 per year of 1 percentage point increase in the share of closeby population choosing the same clock location. This implies, for example, that the social interaction component of the shift premium required to induce somebody working in the 9am shift to work at the 11pm shift (for a reduction of 15.5 percentage points in the share of nearby population choosing the same clock location ) would be approximately 41%[13]

The WTP estimates imply that clock location 23 (11pm) should command a premium of 18+41=59%. Indeed if we re-estimate the simple day work/shift work model as in Kostiuk (1990) but only using workers who start to work at 9am (as day workers) or 11pm (as shift workers) we find that the selection corrected premium is 53%, fairly close to the number derived from our clock location model (59%). The advantage of our approach is that we can tease out how much of this premium is due to something that we can do nothing about (i.e. reduced exposure to daylight) and how much of it is really due to something (i.e. low companionship) that is more malleable and can actually change for the better in alternative equilibria.

## 3.7 Conclusions

Willingness to pay for daylight is shown to be a significant, but by no means exhaustive component of shift premia required to elicit labor supply at non-standard hours. Our estimates account for social interactions externalities in the timing of work and for the endogenous sorting across locations around the clock. They imply that the wage premia required to elicit labor supply at non-standard hours is not

[13]Taking the average hourly wage in the sample \$15.08 we get ($\frac{0.39\times 15.5}{14.6} = 0.41$)

an immutable, hard wired feature of preferences for daylight. Indeed, our estimates imply that only ($\frac{18}{59} \times 100 =$) 30 percent of the premium is due to compensating differential for darkness disamenity. Moreover, from each individual's own perspective, the required premium becomes smaller, the larger the fraction of the population that will be accompanying him/her at those hours. Our results have intriguing implications for potential welfare gains associated with an alternative equilibrium in which the workforce is more evenly distributed around the clock: while we have large productivity gains spurred by more efficient use of capital and infrastructure, welfare loses associated with less exposure to daylight during the night hours are much smaller when a significant fraction of the population is working at those hours. The implication is that one can compensate a large part of the welfare loss induced by reduced exposure to daylight by having a substantial share of the population providing companionship at non-standard hours.

## 3.8 Tables

Table 3.1: Descriptive Statistics

| Variable | Mean | Std. Dev. |
|---|---|---|
| 0 to 5 years of education | 0.01 | 0.11 |
| 6 to 9 years of education | 0.04 | 0.19 |
| 10 to 13 years of education | 0.69 | 0.46 |
| 14 to 16 years of education | 0.23 | 0.42 |
| more than 16 years of education | 0.03 | 0.18 |
| Years of Experience | 21.21 | 9.38 |
| Male | 0.81 | 0.39 |
| Non-White | 0.2 | 0.4 |
| Veteran | 0.19 | 0.39 |
| Single and with at least 1 child under 13 | 0.06 | 0.24 |

Source: 2000 Census of Population. Final Analysis Sample Size N= 13576 corresponding to 16 bundles of PUMAs. See Appendix for details on sample construction.

Table 3.2: OLS and Selection Correction Estimates of Log Wage Equations for Day and Shift Workers

| | OLS | | Selection Correction | |
|---|---|---|---|---|
| Variable | shift | day | shift | day |
| Experience | 0.025*** | 0.044*** | 0.027*** | 0.049*** |
| | (0.009) | (0.005) | (0.009) | (0.006) |
| $Experience^2$ | -0.001 | -0.002*** | -0.001 | -0.002*** |
| | (0) | (0) | (0) | (0) |
| $Experience^3$ | 0.001 | 0.002*** | 0.001 | 0.002*** |
| | (0.001) | (0) | (0.001) | (0) |
| Male | 0.207*** | 0.266*** | 0.201*** | 0.254*** |
| | (0.023) | (0.014) | (0.022) | (0.015) |
| Nonwhite | -0.069*** | -0.047*** | -0.073*** | -0.066*** |
| | (0.021) | (0.013) | (0.02) | (0.014) |
| Married | 0.075*** | 0.091*** | 0.080*** | 0.100*** |
| | (0.017) | (0.01) | (0.017) | (0.011) |
| Veteran | -0.011 | 0.011 | -0.016 | -0.001 |
| | (0.02) | (0.012) | (0.02) | (0.013) |
| SMSA | 0.046** | 0.067*** | 0.049*** | 0.079*** |
| | (0.019) | (0.01) | (0.018) | (0.011) |
| 6 to 9 yrs of education | 0.161* | 0.139*** | 0.160* | 0.139*** |
| | (0.091) | (0.051) | (0.082) | (0.049) |
| 10 to 13 yrs of education | 0.308*** | 0.311*** | 0.312*** | 0.325*** |
| | (0.081) | (0.047) | (0.075) | (0.045) |
| 14 to 16 yrs of education | 0.440*** | 0.408*** | 0.444*** | 0.426*** |
| | (0.082) | (0.048) | (0.077) | (0.046) |
| gt 16 yrs of education | 0.409*** | 0.500*** | 0.413*** | 0.528*** |
| | (0.09) | (0.057) | (0.086) | (0.053) |
| Inverse Mills Ratio | | | -0.103** | 0.382*** |
| | | | (0.041) | (0.039) |
| Constant | 1.931*** | 1.964*** | 2.019*** | 1.740*** |
| | (0.108) | (0.068) | (0.113) | (0.071) |
| Observations | 3444 | 10132 | 3444 | 10132 |

Robust standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. All specifications include (not show in table) spatial bundle effects and occupation effects. Analysis sample based on 2000 Population Census.

Table 3.3: Shift Differentials (% premium over standard day work)

| | | |
|---|---|---|
| All | 4.6 | 39 |
| White | 5 | 39.4 |
| Non-White | 2.7 | 38 |

Shift premia computed based on estimated wage equations using our final analysis sample from the 2000 Population Census.

Table 3.4: Parameters of the Clock Location Model

| Variable | Estimate | Standard Error |
|---|---|---|
| $\alpha_0$ | 9.44 | (1.46) |
| $\alpha_1$ | 0.93 | (0.73) |
| $\beta_0$ | 0.11 | (0.03) |
| $\beta_1$ | -0.04 | (0.09) |
| $\gamma$ | 0.24 | (0.41) |

Table 3.5: Estimates of Willingess to Pay (WTP) for Daylight

| | Hourly | Daily | Weekly | Annual |
|---|---|---|---|---|
| | 2nd Stage: OLS | | | |
| Female | $ 0.21 | $1.71 | $8.57 | $428.58 |
| Male | $ 0.39 | $3.13 | $15.65 | $782.46 |
| | 2nd Stage: IV | | | |
| Female | $ 0.27 | $2.18 | $10.89 | $544.25 |
| Male | $ 0.45 | $3.59 | $17.97 | $898.26 |

Table 3.6: Estimates of Willingess to Pay (WTP) for Companionship

| | Hourly | Daily | Weekly | Annual |
|---|---|---|---|---|
| | | 2nd Stage: OLS | | |
| Female | $ 0.47 | $3.73 | $18.66 | $932.79 |
| Male | $ 0.43 | $3.42 | $17.12 | $855.96 |
| | | 2nd Stage: IV | | |
| Female | $ 0.43 | $3.43 | $17.15 | $857.25 |
| Male | $ 0.39 | $3.12 | $15.61 | $780.44 |

# Appendices

## 3.A Spatial Bundles

### Construction of Spatial Bundles:

We started with the variable defined as Consistent PUMA (CONSPUMA) in the census data. CONSPUMA identifies the most detailed geographic areas that can consistently be identified across samples from 1980 onward. This variable is fully comparable across years. Each CONSPUMA consists of many PUMAs and the data provides the latitude and longitude of each PUMA. We calculate the centroid of each CONSPUMA by averaging the latitude and longitude of each PUMA in that CONSPUMA. We then calculate the distance between all the CONSPUMA's using the Great Circle Distance formula which is the shortest distance between any two points on the surface of a sphere measured along a path on the surface of the sphere.[14] Taking the centroid of each CONSPUMA as the center of a bundle, all CONSPUMAs that were within 10 miles of the centroid were counted as a part of that bundle. The reason for this is that in CONSPUMAs that are close enough to

[14]The formula used is:

$$distance = 6371.01 * arctan\left(\frac{\sqrt{(cos\phi_2 sin\Delta\lambda)^2 + (cos\phi_1 sin\phi_2 - sin\phi_1 cos\phi_2 cos\Delta\lambda)^2}}{sin\phi_1 sin\phi_2 + cos\phi_1 cos\phi_2 cos\Delta\lambda}\right)$$

where,

- $\phi_1, \lambda_1; \phi_2, \lambda_2$ are the geographical latitude and longitude of the two CONSPUMA's
- $\Delta\lambda$ is the difference in longitude between the two CONSPUMA's

each other the percentage of people working in a given shift might be determined by people belonging to all the CONSPUMAs. Hence the percentage in own shift for the people in the CONSPUMA was calculated by including all the people in the bundle

## 3.B Analysis Sample

The final analysis sample used in estimation is constructed in the following way: since we use annual earnings to construct our measure of hourly wage, we keep individuals who worked between 48 and 52 weeks last year. Moreover, we only keep those individuals who report usually working 40 hours per week. We restrict our sample to individuals 18 to 55 year old. We also discard those observations for whom the time they depart from home to work is missing as this prevents us from constructing a measure of location around the clock. Finally to minimize the incidence of outliers we drop observations whose hourly wage is less that one dollar. Finally, we drop all occupations in spatial bundles that have at least one clock location without a single worker. This leads to several small spatial bundles being dropped as no occupation satisfies this requirement.

# Bibliography

**Alesina, Alberto F., Edward L. Glaeser, and Bruce Sacerdote.** 2006. "Work and Leisure in the U.S. and Europe: Why So Different?" In *NBER Macroeconomics Annual 2005, Volume 20.*: National Bureau of Economic Research Inc, 1–100.

**Bayer, Patrick, Robert McMillan, and Kim Rueben.** 2004. "An Equilibrium Model of Sorting in an Urban Housing Market." NBER Working Papers 10865, National Bureau of Economic Research, Inc.

**Bayer, Patrick, and Christopher Timmins.** 2005. "On the equilibrium properties of locational sorting models." *Journal of Urban Economics*, 57(3): 462 – 477.

**Bayer, Patrick, and Christopher Timmins.** 2007. "Estimating Equilibrium Models Of Sorting Across Locations." *The Economic Journal*, 117(518): 353–374.

**Bergstrom, T. C.** 1976. "Toward a Deeper Economics of Sleeping." *Journal of Political Economy*, 84(2): 411–412.

**Berry, Steven, James Levinsohn, and Ariel Pakes.** 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–90.

**Betancourt, R., and C. Clague.** 1981. *Capital Utilization: A Theorethical and Empirical Analysis.* New York: Cambridge University Press.

**Betancourt, Roger R.** 1986. "A Generalization of Modern Production Theory." *Applied Economics*, 18(8): 915–928.

**Biddle, Jeff E., and Daniel S. Hamermesh.** 1990. "Sleep and the Allocation of Time." *Journal of Political Economy*, 98(5): 922–943.

**Calvo, Guillermo A.** 1975. "Efficient and Optimal Utilization of Capital Services." *The American Economic Review*, 65(1): pp. 181–186.

**Dantzig, G. B., and T. L. Saaty.** 1973. *Compact City: Plan for a Liveable Urban Environment.* San Francisco: W. H. Freeman.

**Foss, Murray F.** 1981. "Long-Run Changes in the Workweek of Fixed Capital." *The American Economic Review*, 71(2): 58–63.

**Georgescu-Roegen, Nicholas.** 1970. "The Economics of Production." *American Economic Review*, 60(2): 1–9.

**Hamermesh, Daniel, and Gerard Pfann.** 2005b. *Capital Utilization: A Theorethical and Empirical Analysis.* Contribution in Economics Series, 271: Elsevier.

**Hamermesh, Daniel S.** 1999. "The Timing of Work over Time." *Economic Journal*, 109(452): 37–66.

**Hamermesh, Daniel S.** 2002. "Timing, Togetherness and Time Windfalls." *Journal of Population Economics*, 15(4): 601–623.

**Hamermesh, Daniel S., Caitlin Knowles Myers, and Mark L. Pocock.** 2008. "Cues for Timing and Coordination: Latitude, Letterman, and Longitude." *Journal of Labor Economics*, 26(2): 223–246.

**Hamermesh, Daniel S., and Gerard A. Pfann.** 2005a. "Time Use Data in Economics." *European Economic Review*, 49(1): 1–7.

**Hoffman, Emily P.** 1977. "The Deeper Economics of Sleeping: Important Clues toward the Discovery of Activity X." *Journal of Political Economy*, 85(3): 647–650.

**Hwang, Hae-shin, W. Robert Reed, and Carlton Hubbard.** 1992. "Compensating Wage Differentials and Unobserved Productivity." *Journal of Political Economy*, 100(4): pp. 835–858.

**Jenkins, Stephen, and Lars Osberg.** 2004. "Nobody to Play With? The Implications of Leisure Coordination." In *The Economics of Time Use: Contributions to Economic Analysis (CEA)*. eds. by Daniel Hamermesh, and Gerard Pfann: Elsevier, Amsterdam, The Netherlands.

**Kahn, Matthew E.** 2008. "Compensating Differentials." In *The New Palgrave Dictionary of Economics, Second Edition*. eds. by Steven N. Durlauf, and Lawrence E. Blume: Palgrave Macmillan.

**Kostiuk, Peter F.** 1990. "Compensating Differentials for Shift Work." *Journal of Political Economy*, 98(5): 1054–75.

**Marris, R.** 1964. *The Economics of Capital Utilization. A Report on Multiple Shift Work*. Cambridge: Cambridge University Press.

**Presser, Harriet.** 2004. *Working in a 24/7 Economy. Challenges for American Families*. New York: Russell Sage Foundation.

**Rosen, Sherwin.** 1987. "The Theory of Equalizing Differences." In *Handbook of Labor Economics*. eds. by O. Ashenfelter, and R. Layard, 1 of Handbook of Labor Economics: Elsevier, , Chap. 12 641–692.

**Weiss, Yoram.** 1996. "Synchronization of Work Schedules." *International Economic Review*, 37(1): 157–79.

**Winston, Gordon C.** 1974. "The Theory of Capital Utilization and Idleness." *Journal of Economic Literature*, 12(4): pp. 1301–1320.

**Yaniv, Gideon.** 2004. "Insomnia, Biological Clock, and the Bedtime Decision: an Economic Perspective." *Health Economics*, 13(1): 1–8.