Living Together – Essays on Cohabitation

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Kelvin Kai Wing Wong

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

José Victor Rios Rull and Larry Jones

July, 2016

Acknowledgements

I thank José Victor Rios Rull for his patience, support, and guidance, Enoch Hill, Kai Ding, Jack Rossbach, Elena Pastorino, and Thomas Holmes for their exceptional support and excellent comments, and Larry Jones, Rishabh Kirpalani, Filippo Rebessi, and Brian Zurowski for their many invaluable suggestions and feedback. A special thanks to Rob Warren for his very involved help in securing the restricted geocoded data used in the second essay.

Dedication

This thesis is dedicated to my wife Angela, for her unending love, prayers, encouragement, and support.

Abstract

This thesis studies cohabitation behavior in the United States, and proposes two answers as to why cohabitation rate increased in the last four decades. In Chapter 1, I propose that the increasing value of economies of scale can contribute to this increase in cohabitation. In Chapter 2, I propose that declining migration could have been a factor.

In Chapter 1, I use the National Longitudinal Survey of Youth 1997 to document that even from before the start of cohabitation, cohabitors who eventually marry their partners and cohabitors who do not have different labor supply and wages. I find that those who cohabit and eventually marry work on average six hours more per week than those who cohabit but will not marry. I estimate that every hour worked per week in the year prior to the start of cohabitation increases the probability of transition into marriage by .7%. Additionally, I find that lower wages prior to the start of cohabitation are associated with a lower probability of transitioning into marriage. I develop a theory of co-residential relationship formation where lower wages increase the value of living together, leading to less selectivity in match quality and thus a lower transition probability into marriage. Singles who have higher hours worked have higher disutility to begin a relationship along with higher income, leading to higher selectivity in match quality and thus a higher probability of transitioning into marriage. I show that lower real wages in the last five decades can explain the trends in cohabitation and marriage for college and non-college graduates. Finally, I suggest that changes to the incentives to form a co-residential relationship create an additional channel where urban poverty programs can affect welfare.

In Chapter 2, I compare the migration and cohabitation behaviors between the National Longitudinal Survey of Youth 1979 and 1997 cohorts. I document that for both cohorts, college graduates migrate more and cohabit less than non-college graduates. I hypothesize that having a higher probability of a future migration decreases the likelihood of cohabiting, as cohabitation can be seen as a form of settling down. I find that college graduates are more likely to move out of county or state. I also find that having a big move in the future decreases the odds of cohabiting. This can explain the

cross-sectional difference between non-college and college graduates cohabitation rate. Comparing across cohorts and using the Census Bureau Current Population Survey, I find that cohabitation has increased and migration has decreased. This implies declining migration could also have contributed to the increase in cohabitation.

Contents

List of Tables

List of Figures

Chapter 1

Settling Down or Settling In?

1.1 Introduction

Marriage and cohabitation[1] differently influence a broad set of economic decisions; from labor supply, to fertility and investment in children, to household production. Hence, the transition from cohabitation to marriage changes how individuals behave. But are some cohabitors more likely to make this transition? Empirical evidence from the National Longitudinal Survey of Youth 1997 (NLSY97) shows that cohabitors who eventually marry are different even before the start of cohabitation than cohabitors who do not marry. In this paper, I analyze how an individual's labor market decisions and characteristics prior to cohabitation influence his or her transition from cohabitation into marriage.

I use the NLSY97, a monthly panel that had previously not been used in the literature, to track individuals in co-residential relationships. The monthly frequency of the data is crucial since cohabitation is often short-lived[2]. Additionally, cross-sectional, annual, or bi-annual data do not as accurately represent labor market characteristics and decisions at or before relationship formation. I first document from the data that college graduates, compared to non-college graduates, cohabit less and marry more. Second, within education groups, cohabitors who eventually marry work on average

[1] Cohabitation in the data is self-reported, and is defined as an unmarried partnership of two opposite-sex individuals who share the same household.

[2] Half of cohabitation in the NLSY97 ends within 12 months.

17% more[3] before the start of cohabitation than cohabitors who do not marry[4]. I estimate that every hour worked in the year prior to the start of cohabitation increases the transition into marriage by about .7%. Lastly, within education groups, those with a non-traditional work schedule[5] prior to cohabitation are less likely to transition into marriage from cohabitation.

To give an intuition as to how labor market characteristics and decisions before even meeting a potential partner could have an impact on the transition of cohabitation into marriage, I develop a theory of relationship[6] formation. Cohabitation and marriage both offer two benefits: economies of scale and relational utility, but require an initial effort to begin, which brings disutility. I use data on dating in the NLSY97 to relate to effort. In my mechanism, those who choose to cohabit have a probability of transitioning into marriage in the future, but this transition requires additional effort. While cohabitation requires less initial effort than marriage, a cohabiting relationship that transitions into marriage requires a higher total effort than a marriage not preceded by cohabitation. This is the fundamental trade-off between cohabiting or not cohabiting before marriage in my model. I show that an individual will marry without first cohabiting only if initial match quality is high[7]. Transitioning into marriage after cohabitation is not an endogenous choice. The probability of cohabitation transitioning into marriage is positively correlated with initial match quality. Hence, the main channel through which labor market characteristics and decisions prior to the start of cohabitation affect the transition probability into marriage is through an individual's initial selectivity over match quality.

For instance, singles with lower wages have a lower value of being single, which implies there is more to gain from the economies of scale aspect of a co-residential

[3] This translates to five to six hours per week.

[4] As can be seen in the data section, this is true for non-college and college graduated men and non-college graduated women, but the difference is not significant for college-graduated women.

[5] Non-traditional work schedules include jobs with night, evening, or irregular work hours.

[6] Any use of the term "relationship" in this paper should be understood as "co-residential relationship"

[7] In this paper, my main focus is not on the difference between marriage and cohabitation, but rather the transition from cohabitation into marriage influenced by pre-relationship factors. Having a model where marriage without cohabiting first is only chosen because of high match quality simplifies the problem. Proposition 1 shows that the decision to marry is not based on labor market characteristics or decisions.

relationship. A larger benefit from economies of scale leads to a wider range of potential match quality that a single individual is willing to initially accept. Thus, individuals with low wages are less selective with whom they cohabit with. The lower average match quality of cohabitors with low wages leads to a lower probability of that their cohabiting relationships will transition into marriage. Since non-college graduates and those with non-traditional work schedules have lower wages, they are more likely to cohabit without eventually marrying.

The same mechanism also explains why individuals who work more before cohabitation have a higher probability of transitioning into marriage. The more hours an individual works, the greater the disutility from effort to start a relationship. In other words, working more increases the cost of starting a relationship. Working more also increases income, which causes the value of economies of scale to fall. Therefore, those with high hours worked are more selective over the match quality of their potential partners. Those that do choose to cohabit will thus have a higher average match quality, which leads to a higher probability of transitioning into marriage. I document using county-level data that increases in the cost of housing is correlated with an increase in the proportion in a cohabiting relationship relative to the proportion of singles. I also show that for non-college graduates, an increase in the cost of housing does not have a significant impact on the proportion married relative to the proportion of singles. This result does not hold for college-graduates, suggesting my theory is better suited to explain cohabitation rates among low income individuals.

My mechanism suggests that trends in cohabitation and marriage in the last five decades could partially be explained by changing real wages. Over the last five decades, cohabitation rates have increased for both college and non-college graduates, but the increase is greater for non-college graduates. Marriage rates have also fallen in the same period, once again more for non-college graduates than college graduates. My mechanism shows that more cohabitation and less marriage will be the result of a decline in real wages over recent decades, which Kambourov and Manovskii (2009) documents. Kambourov and Manovskii also documents that there had been a flattening of life cycle real annual earnings that is greater for the non-college graduates than college graduates. Lower real wages lead individuals to cohabit more because the value of economies of scale in the household is larger, leading cohabitation rates to increase. Importantly,

the average match quality of these cohabiting relationships are lower because of the increasing value of economies of scale, leading to a lower probability of transition into marriage. Since 62% of marriages today begin from cohabitation, this implies that the marriage rate will also fall. While changes in the trends of marriage and cohabitation are usually attributed to changes in culture and societal norms, I suggest that changes in real wages over this time period can also contribute to these trends. Lastly, I suggest that welfare analysis of urban poverty programs should account for changes in incentives to form co-residential relationships. Programs that aim to increase wages, whether through education subsidies or by providing better access to jobs with traditional work schedules, decrease the value of economies of scale. As a result, individuals are more selective over match quality and cohabiting relationships with low match qualities are less likely to begin. I suggest that this has implications on welfare through impacts on the occurrences of domestic violence, fertility, and investment in children.

1.2 Related Literature

Becker (1973) and Becker (1974) are commonly referred to as the first to apply economic theory to analyze marriage, though his analysis is more broadly applicable to co-residential relationships. He proposes a model of household production that depends on the number of individuals in a household, and shows that whether marriage occurs or not depends on the male and female's income, human capital, and relative difference in wage rates. Becker (1981) formalizes these ideas and proposes more theories of relationship formation based on sorting.

As cohabitation became increasingly prevalent in recent decades, so had research that focused on cohabitation's differences from marriage in regards to formation, fertility, child investment, and impacts on labor market behavior[8]. Lillard et al. (1995) proposes that individuals who are prone to divorce if they married are the same ones who are likely to select into cohabitation. This selection explains why those who cohabit

[8] Smock and Manning (2004) suggests that the interest in understanding cohabitation stems from its linkage to the decreasing centrality of marriage in the United States. Additionally, they summarize the literature to conclude that cohabitation is seen in three ways: as a stage in the marriage process, an alternative to marriage, or an alternative to singlehood. The theory presented in this paper is general enough such that all three interpretations of cohabitation is possible.

before marriage tend to have higher divorce rates. Brien et al. (2006) models cohabitation as a way of learning about a relationship's true match quality, after observing a noisy signal when the match first takes place. Adamopoulou (2010) analyzes wage gap and improvement in household production technology on marital and cohabitation decisions of Europeans. Greenwood et al. (2012) proposes that technological progress in the household sector and shifts in wage structure have changed incentives to marry. Lundberg and Pollak (2013) suggests that co-residential relationships are a way of choosing an investment strategy for children, where marriage is preferred if there will be high investment in one's children. Gemici and Laufer (2014) presents a model of comparative advantage between market and house work. When gains to specialization are high, it is optimal for the woman to specialize in home production. However, women give up on acquiring human capital by specializing in home production, and thus must be compensated by a lower probability of relationship separation through marriage. Cohabitors are thus those who do not have high gains in specialization.

Other papers have focused on the impact of policy on relationship formation. Light and Omori (2008) shows that different policies influence union formation. They find that an increase in state income tax have a small and statistically insignificant effect on the transition of white and Hispanic females into marriage (either from single or cohabiting), but a significant decrease of black cohabiting women transitioning into marriage. Chade and Ventura (2005) analyzes the impact of different tax treatments between the single and married, and finds that while a tax penalty for being married reduces the number of marriages, it does not necessarily make men and women more hesitant to marry. In an extension where cohabitation is added, they find that the number of marriages becomes more sensitive to a marriage tax penalty as individuals who were close to the threshold between being single and marrying will now choose to cohabit. Blau and van der Klaauw (2010) considers how changes in policies, divorce and tax laws, and wages jointly affect the family structure that a child will grow up in.

My departure from the literature is three-fold. First, I look at how individual characteristics before the start of cohabitation influence the transition probability into marriage. Second, this is the first paper to use a monthly panel to more accurately analyze relationship formation, which is important because cross-sectional or annual/bi-annual data will not accurately provide labor market characteristics at the time of relationship

formation. Third, relationship formation is not assumed to be costless because of an initial effort that must be exerted before the start of a co-residential relationship[9].

The paper proceeds as follows. Section 2 shows data from the NLSY97, Section 3 presents the model, Section 4 shows analytical solutions that matches with data, Section 5 applies the model to analyze trends in cohabitation and marriage rates and evaluate policy impact, and Section 6 concludes.

1.3 Data

In this section, I describe the NLSY97 data and present three main empirical findings: 1) College graduates, compared to non-college graduates, are less likely to cohabit and more likely to marry (with or without cohabiting first), 2) those who cohabit and later marry on average work more than those who cohabit but do not marry, 3) those who have non-traditional work schedules are more likely to cohabit without marrying.

1.3.1 Data Description

The National Longitudinal Survey of Youth 1997 (NLSY97) is a monthly panel that tracks 8984 youths born between 1980 and 1984[10]. The NLSY97 is a relatively unexplored dataset because the oldest respondents, as of Round 15, are only 32 years old. Thus, this paper will focus on adults between ages 21 and 32, from which there are 6468 co-residential relationships (marriages or cohabiting relationships).

Cohabitation and marriage are self-reported in the data. The respondents are asked about their marital or cohabitation status for each month in the previous year within the same question, and thus cohabitation is implied to mean more than just two people living in the same place. Unique identifiers are given to each spouse or partner, allowing each relationship to be tracked over time. I construct a relationship timeline for each relationship. Details can be found in the appendix.

[9] Chade and Ventura (2005) and Smith (2006) use models of search for potential partners with friction (time cost), but to my knowledge this is not used outside of the search literature. I extend on this idea by providing empirical evidence and using effort in a model without search.

[10] Rounds 1 to 15 of the NLSY97 were used.

1.3.2 Types of Relationships by Education

I categorize relationships into three different groups. A relationship is categorized as "cohabit" if the respondent has a cohabiting partner but marriage with that partner is not observed. This could mean either I observe an end to cohabitation without marriage, or the relationship is right censored and cohabitation was the last observed relationship status[11]. A relationship is categorized as "marry" if the respondent is not observed to have lived with their partner before marriage. However, there is a possibility that cohabitation of less than one month occurred before marriage and is not indicated in the data. Thus, the "marriage" group may include those who cohabited for less than one month before marriage. Finally, a relationship is classified as "cohabit and marry" if the respondent cohabited with a partner and then eventually marries that partner.

A summary of the number of relationships in each relationship group, further divided by education at the start of the relationship, is presented in Table 1.1. Details can be found in the appendix on how education groups were constructed for each respondent-month, along with more descriptive statistics based on education.

Table 1.1: Number of Relationships by Relationship Groups and Education Level

	Non-College Graduate	College Graduate	Total
Cohabit	3022 (55.3%)	375 (37.6%)	3397
Cohabit and Marry	1533 (28.0%)	392 (39.2%)	1925
Marry	913 (16.7%)	233 (23.3%)	1146
Total	5468 (100%)	1000 (100%)	6468

[11] Robustness checks in the appendix shows that it does not matter whether I further classify the cohabit only group into "cohabit and end to relationship was observed" and "cohabit and end to relationship was not observed". Taking out cohabitations that are right censored will affect college graduates more than non-college graduates, since college graduates on average start relationships later than non-college graduates. This is also shown in the appendix.

Four main observations from Table 1.1:

1. College graduates cohabit less than non-college graduates ($37.6\% < 52.3\%$)
2. College graduates marry from cohabitation more than non-college graduates ($39.2\% > 28.0\%$)
3. College graduates marry without cohabiting first more than non-college graduates ($23.3\% > 16.7\%$)
4. For both non-college and college graduates, about 63% of marriages begin from cohabitation

1.3.3 Income

Average weekly income at the start of a co-residential relationship, in 2005 dollars, is shown in Table 1.2 for different education and age groups.

Table 1.2: Average Weekly Income in 2005 dollars, at Start of Relationship

	Age 21-24	Age 25-32
Non-College	\$442.06	\$536.04
College	\$608.75	\$786.13

College graduates have higher weekly income than non-college graduates within each age group. The result is the same if I use age instead of age group. Details on how weekly income was calculated can be found in the appendix.

1.3.4 Hours Worked

For those in the cohabit and marry group, labor hours supplied are higher at the start of a relationship than those who only cohabit or only marry. The result of a regression of weekly hours worked (at the beginning of a relationship) on education interacted with relationship group and other control variables is shown in Table 1.3.

The coefficient for each education-relationship group in Table 1.3 compares weekly hours worked with non-college graduates who only cohabit (the reference group). Within each education group, those in the cohabit and marry group will work more than those

who only cohabit or only marry. For example, hours worked at the start of the relationship is 1.75 hours more per week for non-college graduate cohabitors who are observed to marry their cohabiting partners at some point in the future than non-college graduate cohabitors who are not observed to marry. For college graduate cohabitors, this difference is 2.75 hours per week.

Table 1.3: Weekly Hours Worked at Start of Relationship

	Coefficient	(Std. Error)
Reference: Non-college & Cohabit		
Non-college & Cohabit and Marry	1.75***	(0.55)
Non-college & Marry only	1.35**	(0.65)
College & Cohabit only	3.03***	(0.89)
College & Cohabit and Marry	5.78***	(0.89)
College & Marry only	4.48***	(1.09)
Female	-3.95***	(0.43)
Age	4.85***	(1.81)
Age Squared	-0.09**	(0.04)
One Child	-0.40	(1.42)
Two Children	2.55	(2.21)
More than Two Children	1.57	(4.12)
Real Hourly Wage [12]	-0.07***	(.01)
Intercept	-23.19	(22.31)

*** denotes statistical significance at a 1% level and ** denotes significance at the 5% level.

This difference also exists prior to the start of the relationship. Figure 1 shows weekly hours worked for male and female college and non-college graduates without children broken into the three relationship groups. Time 0 is the start of the relationship, and thus the figure shows average weekly hours worked from 12 months before a relationship

[12] Real hourly wage is the average real hourly wage of an individual in any given week across concurrent jobs.

(conditional on respondent being single) to the start of the relationship. Those who cohabit and eventually marry work more prior to the start of a relationship[13]. In other words, even before cohabiting with their partners, those who will eventually marry from cohabitation work more each week than those who cohabit but do not marry or those who marry without cohabiting. On average, the difference is about 6 hours more per week, or 17% more hours worked, for those who cohabit and marry than those who only cohabit. This difference between the two groups in hours worked is lowest for college educated women.

The appendix provides details on relationship time construction, how hours worked is constructed from variables in the NLSY97, and descriptive statistics of labor force participation and employment.

I use a probit regression to determine the probability increase in marrying conditional on cohabitation from each additional hours worked per week. Table 1.4 shows the results of the regression. The sample of this regression is restricted to those who cohabit, whether they marry eventually or not. The dependent variable is equal to 0 if marriage does not occur from the cohabitation, and equal to 1 if marriage occurs. The main variable of interest is how an additional hour worked each week will change the probability of marriage from cohabitation. Table 1.4 shows that even 12 months prior

[13] Those who are married generally work more than those who cohabit, and those who cohabit and eventually marry generally work more than those who marry. While not shown in Figure 1, those who cohabit and eventually marry continue to work more after the relationship begins than those who only cohabit. Of particular interest in the literature is how female hours worked change after cohabitation or marriage. Gemici and Laufer (2014) observes that hours worked by women who marry are lower than hours worked by women who cohabit, and suggests that after marriage, women's labor market hours should fall because of specialization in household production. However, I do not observe this in the data for non-college females. Those who cohabit and later marry persist in working more than cohabitors even after marriage. Meaning that a subset of married women (those who first cohabited) actually work more than those who are cohabiting. This is less clear for college-educated female, as market hours worked does not seem to change significantly after marriage or cohabitation. While a careful analysis of the response of hours worked to relationship formation is not conducted in this paper, an initial look at the data suggests that it is important to classify those who marry from cohabitation differently than those who marry without cohabiting first. Lastly, one reason as to why I arrive at a different result as Gemici and Laufer (2014) could be because I use the NLSY97 instead of the NLSY79. It is obvious that much had changed in the two decades between the NLSY79 cohort and the NLSY97 cohort.

to the start of a cohabitation, the probability of marriage will increase by .7% for each additional hour worked each week. To put this into perspective, suppose two individuals will cohabit 12 months from today, and both have similar characteristics except one works 40 hours a week and the other works 20 hours a week. It is predicted that the individual who works 40 hours a week will have a 14% higher probability of marrying from cohabitation than the individual who works 20 hours a week twelve months, even though the cohabitation will not begin for another twelve months.

Figure 1.1: Weekly Hours Worked Before the Start of Relationship

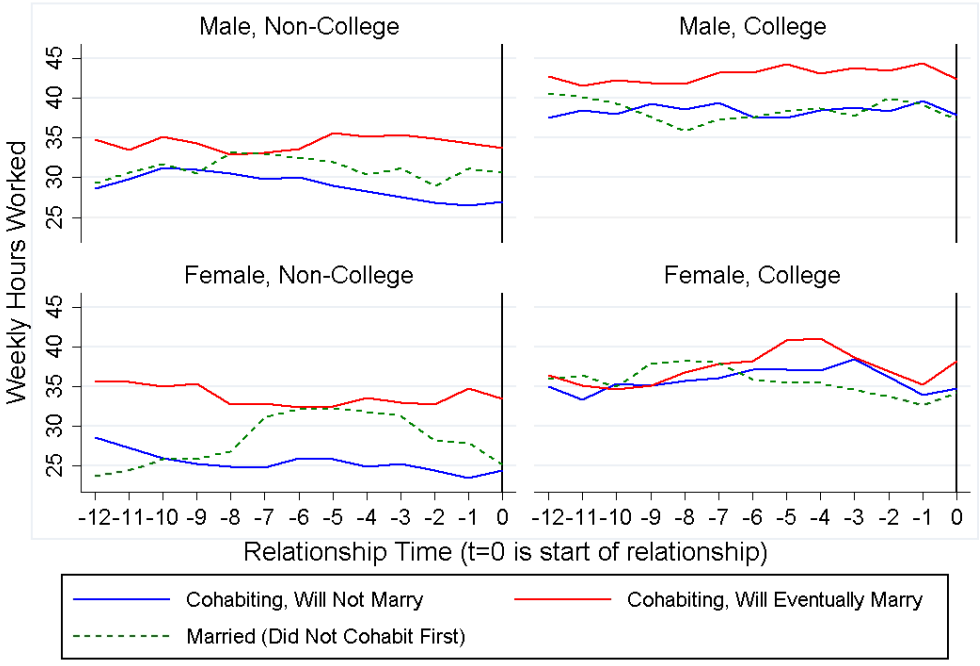


Table 1.4: Probability of Marriage after Cohabitation

	t=0		t=-3		t=-6		t=-12	
Hours worked(wk)	.007***	(.002)	.006***	(.002)	.006***	(.002)	.007***	(.002)
Female	-.002	(.05)	-.001	(.05)	-.02	(.06)	-.06	(.06)
College Degree	.46***	(.07)	.53***	(.07)	.50***	(.08)	.42***	(.08)
Real Wage	.002	(.001)	.001	(.001)	.0009	(.001)	-.002	(.002)
Age	.10***	(.22)	.41*	(.24)	.70***	(.26)	.46*	(.27)
Age Squared	-.003***	(.004)	-.01**	(.005)	-.02***	(.005)	-.01**	(.006)
One Child	.02	(.16)	-.48*	(.27)	-.20	(.30)	— [14]	—
Two Children	-.34	(.30)	.25	(.46)	-.63	(.57)	.06	(.35)
>Two Children	.84	(.45)	-.02	(.69)	.59	(.89)	.28	(.54)
Intercept	-1.53	(2.72)	-5.06*	(2.96)	-8.40***	(3.15)	-5.15	(3.33)

Numbers given are coefficients to probit regression, and in parenthesis are standard errors. *** denotes statistical significance at a 1% level, ** denotes significance at the 5% level, * denotes significance at the 10% level.

1.3.5 Job Schedule

A traditional work schedule is one with a regular daytime schedule, whereas a non-traditional work schedule is one with evening, night, or irregular shifts. The average hourly wage of those with a non-traditional work schedule in their primary job (the job with the most hours worked each month) is lower than those with a traditional work schedule in their primary job, as shown in Table 1.5[15]. Thus, a non-traditional work schedule is correlated with lower wages. Because this is a within group comparison, college graduates should not be compared with non-college graduates. Table 1.6 shows that non-college graduates are more likely to have a non-traditional job schedule.

[14] Omitted because of the small sample size of those with one child twelve months before the start of a cohabitation.

[15] Wages are in 2005 dollars.

Table 1.5: Average Real Hourly Wage of Individual by Job Schedule of Primary Job

	Non-College		College	
	Age 21-24	Age 25-32	Age 21-24	Age 25-32
Favorable	$12.03	$13.36	$16.01	$19.02
Non-traditional	$10.39	$12.33	$12.30	$17.04
Difference (%)	-13.6%	-7.7%	-23.2%	-10.4%

Table 1.6: Percentage of Education Group with Each Schedule Type

	Non-College Graduates	College Graduates
Traditional	57.6%	81.8%
Non-traditional	42.4%	18.2%

Non-college graduates between ages 21 and 32 with non-traditional schedules are more likely to cohabit than those with traditional schedules, as shown in Table 1.7. Table 1.8 shows that this is also true for college graduates.

Table 1.7: Relationship Types of Those with Non-traditional Schedules: Non-college

	Traditional	Non-Traditional
Cohabit	49.0%	58.4%
Cohabit and Marry	34.9%	26.6%
Marry	16.0%	15.0%

Table 1.8: Relationship Types of Those with Non-traditional Schedules: College

	Traditional	Non-Traditional
Cohabit	32.2%	52.4%
Cohabit and Marry	43.4%	33.3%
Marry	24.4%	14.29%

1.3.6 Number of Dates

Initial effort of a co-residential relationship is estimated using the number of dates in the period prior to cohabitation or marriage. This is meant to be a measure of the parts of dating that will require effort and brings disutility to the individual. For example, getting to know a potential partner's friends and family requires effort, but is done for the sake of advancing the relationship. More details on how the NLSY97 variable for annual number of dates gone on is used to get the number of dates gone on before the start of a relationship is found in the appendix.

Table 1.9 shows the number of dates before the start of a relationship, by relationship group. Dating before cohabitation is roughly the same regardless of whether marriage happens in the future, and those who marry without cohabiting first go on more dates than those who cohabit. This represents the greater effort required for marriage compared with cohabitation.

Table 1.9: Number of Dates Before Start of Relationship, by Relationship Group

	Average number of dates before start of relationship
Cohabit	27.32
Cohabit and Marry	26.94
Marry	40.11

Table 1.10 estimates how much work schedule and hours worked impact dating. Every hour worked per month with a traditional schedule decreases number of dates by .0066%, and decreases by .0095% for every hour worked with a non-traditional schedule. To put these numbers into perspective, someone who works 30 hours with a traditional schedule will go on 8.6% less dates than someone who does not work, and someone with a non-traditional schedule will go on 12.4% less dates than someone who does not work. At 50 hours, that becomes 14.3% less dates for those with traditional schedules and 20.6% less dates for those with non-traditional schedules.

Table 1.10: Regression on Log Average Number of Dates Per Month

	Coefficient	(Std. Error)
Non-college traditional Schedule (per hour worked)	-0.00065***	(.00008)
Non-college non-traditional Schedule (per hour worked)	-0.00095***	(.00009)
Real Hourly Wage	.0005	(.0003)
Age	.075***	(0.29)
Age Squared	-0.001*	(0.0007)
Cohabiting	-0.09***	(0.02)
Married	-0.69***	(0.07)
Divorced, Legally Separated, or Widowed	-0.38***	(0.08)
College Degree [16]	0.28***	(.036)
Intercept	-1.13***	(.32)

*** denotes statistical significance at a 1% level, ** denotes significance at the 5% level, and * denotes significant at the 10% level. Controlled for individual fixed effects.

1.4 Model

I introduce a simple qualitative model that is consistent with the data presented to give intuition into how labor characteristics and decisions affect relationship formation.

The model has two periods and two types of agents, college and non-college graduates. College and non-college graduates differ in their wages, determined exogenously. Additionally, work schedule, number of hours worked, wages, partner's income, and

[16] An assumption in this paper is that given a college and non-college graduate with the same number of hours worked and work schedule, they will face the same effort to start a relationship. Data is clear that college graduates date more than non-college graduates, though this does not necessarily mean that college graduates face more disutility in starting a relationship than non-college graduates. The data appendix shows more details on dating by education group, and that within each education group, it is still observed that 1) roughly the same effort is seen before the start of a cohabitation, regardless of whether that cohabitation transitions into a marriage in the future and 2) marriage requires more effort than cohabitation.

match quality are exogenous and known prior to the start of period 1. Given the exogenous variables, an individual chooses to stay single, cohabit, or marry in period 1. There is uncertainty through a shock to match quality in period 2. If match quality after the shock is above zero in period 2, and the individual cohabited in period 1, he becomes married. If he married in period 1, he remains married in period 2 only if match quality after the shock is above zero. If the shock brings match quality below zero, the relationship (either cohabitation or marriage) ends. Period 1, then, can be thought of as the five to seven year period where most cohabiting couples either transition into marriage or separate, and where divorce has the highest probability of happening. Period 2 is the time following that period, when most transitions in or out of a relationship have already taken place.

1.4.1 Preferences

Utility for an agent depends on household income and match quality.

$$u = \begin{cases} \ln(w_i(j_x)h) & \text{if single} \\ \ln(w_i(j_x)h + I^\star) + \theta & \text{if in co-residential relationship} \end{cases}$$

Where $w_i(j_x)$ is the wage for $i = \{\text{c}, \text{nc}\}$ (college and non-college) that is dependent on job schedule j_x, with $x = \{f, uf\}$ (traditional and non-traditional)[17], h is hours worked, $I^\star$ is the income of partner, and θ is match quality.

Wages are the only thing that differs between a college graduate and a non-college graduate. It is assumed that college and non-college graduates do not draw different hours worked[18], partner's income, or match quality.

[17] To avoid confusion of subscripts, I use f and uf to denote traditional (can think of as "favorable") and non-traditional (can think of as "unfavorable") schedules.

[18] While in the data, college graduates work more than non-college graduates, allowing for college graduates to have higher hours worked will not change the results. This is because higher hours worked has the same effect on match selectivity as higher wages, and college graduates by definition have higher wages than non-college graduates. Because results are qualitative, not including a difference in hours worked will not matter for results, and will simplify notation.

1.4.2 Household Income

Household income consists of income of an agent plus the income of his partner. The agent's wage can take on four values, $w_{col}(j_f)$, $w_{col}(j_{uf})$, $w_{ncol}(j_f)$, and $w_{ncol}(j_{uf})$, with $w_{col}(j_f) > w_{col}(j_{uf})$ and $w_{ncol}(j_f) > w_{ncol}(j_{uf})$. No comparison is made across education groups for different job schedules – therefore, it does not matter for this model whether the wage of a college graduate with a non-traditional schedule is higher than the wage of a non-college graduate with a traditional schedule. Jobs with traditional work schedules tend to pay more, and college graduates receive higher wages than non-college graduates given the same job schedule, $w_{col}(j_x) > w_{ncol}(j_x)$.

An individual's wage, hours worked, and partner's income are all exogenously determined before the beginning of the first period. In other words, when an agent makes a decision for whether to remain single, cohabit, or marry, he knows his own hours worked, work schedule, wage, and partner's income.

1.4.3 Match Quality

At the start of period 1, agents receive match quality θ of a potential partner drawn randomly from a distribution. In period 2, there is a shock ε to match quality. If $\theta - \varepsilon > 0$, then the relationship survives the shock. This means that a cohabitation will transition into a marriage, or a marriage will continue. For both types of co-residential relationship, if match quality is below zero in period 2, the relationship ends.

1.4.4 Effort

After receiving match quality, an agent can choose to pay effort $e(h, j_x)$ to enter a co-residential relationship. Effort depends on hours worked, h, such that $e(h_{low}, j_x) < e(h_{high}, j_x)$, and whether an agent's work schedule is traditional or non-traditional,$j_x > 0$, where $x = \{f, uf\}$ and $e(h, j_f) < e(h, j_{uf})$. The effort function is defined as follows

$$e(h, j_x) = j_x h^{\alpha} \text{where} 0 < \alpha < 1$$

Marriage is more costly than cohabitation to begin, and agents who wish to marry must pay $pe(h, j_x) = pj_x h^{\alpha}$ where $1 < p < \frac{1}{1-\beta}$.

Agents who cohabit in the first period and marry in the second period pays $j_x h^\alpha$ in the first period and $pj_x h^\alpha$ in the second period. Agents who marry in the first period only pay $pj_x h^\alpha$ in the first period. Therefore, marriage that begins from cohabitation requires more effort over the two periods. The trade-off of cohabiting before marriage is paying less effort in the first period, but more effort in total.

1.4.5 Agent's Problem

The agent solves the following problem in period 1 given $(w_i(j_x), h, j_x, \theta, I^\star)$:

$$\max\{\text{v(single)}, \text{v(cohabit)}, \text{v(marry)}\}$$

where:

$$\begin{aligned}
\text{v(single)} &= \ln(w_i(j_x)h) + \beta \ln(w_i(j_x)h) \\
\text{v(cohab)} &= -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta \\
&\quad + \beta \Bigg[\int_{\underline{\varepsilon}}^{\theta} (\ln(w_i(j_x)h + I^\star) - pj_x h^\alpha + \theta - \varepsilon) f(\varepsilon | \varepsilon \leqslant \theta) d\varepsilon \\
&\quad + \int_{\theta}^{\bar{\varepsilon}} \ln(w_i(j_x)h) f(\varepsilon | \varepsilon > \theta) d\varepsilon \Bigg] \\
\text{v(marry)} &= -pj_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta \\
&\quad + \beta \Bigg[\int_{\underline{\varepsilon}}^{\theta} (\ln(w_i(j_x)h + I^\star) + \theta - \varepsilon) f(\varepsilon | \varepsilon \leqslant \theta) d\varepsilon \\
&\quad + \int_{\theta}^{\bar{\varepsilon}} \ln(w_i(j_x)h) f(\varepsilon | \varepsilon > \theta) d\varepsilon \Bigg]
\end{aligned}$$

Assuming $\varepsilon \sim U[\underline{\varepsilon}, \bar{\varepsilon}]$, and since separation occurs when match quality falls below zero in the second period, the problem can be written as follows:

$$\text{v(single)} = \ln(w_i(j_x)h) + \beta \ln(w_i(j_x)h)$$

$$\text{v(cohab)} = \begin{cases} -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \ln(w_i(j_x)h) & \frac{\theta - \underline{\varepsilon}}{\bar{\varepsilon} - \underline{\varepsilon}} < 0 \\ -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta & \\ \quad + \beta \left[\left(\dfrac{\theta - \underline{\varepsilon}}{\bar{\varepsilon} - \underline{\varepsilon}} \right) \left(\ln(w_i(j_x)h + I^\star) - p j_x h^\alpha + \dfrac{1}{2}\theta - \dfrac{1}{2}\underline{\varepsilon} \right) \right. & 0 \leqslant \frac{\theta - \underline{\varepsilon}}{\bar{\varepsilon} - \underline{\varepsilon}} \leqslant 1 \\ \quad \left. + \left(\dfrac{\bar{\varepsilon} - \theta}{\bar{\varepsilon} - \underline{\varepsilon}} \right) \ln(w_i(j_x)h) \right] & \\ -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta & \\ \quad + \beta \left[\ln(w_i(j_x)h + I^\star) + \dfrac{1}{2}\theta - p j_x h^\alpha \right] & \frac{\theta - \underline{\varepsilon}}{\bar{\varepsilon} - \underline{\varepsilon}} > 1 \end{cases}$$

$$\text{v(marry)} = \begin{cases} -p j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \ln(w_i(j_x)h) & \frac{\theta - \underline{\varepsilon}}{\bar{\varepsilon} - \underline{\varepsilon}} < 0 \\ -p j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta & \\ \quad + \beta \left[\left(\dfrac{\theta - \underline{\varepsilon}}{\bar{\varepsilon} - \underline{\varepsilon}} \right) \left(\ln(w_i(j_x)h + I^\star) + \dfrac{1}{2}\theta - \dfrac{1}{2}\underline{\varepsilon} \right) \right. & 0 \leqslant \frac{\theta - \underline{\varepsilon}}{\bar{\varepsilon} - \underline{\varepsilon}} \leqslant 1 \\ \quad \left. + \left(\dfrac{\bar{\varepsilon} - \theta}{\bar{\varepsilon} - \underline{\varepsilon}} \right) \ln(w_i(j_x)h) \right] & \\ -p j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta & \\ \quad + \beta \left[\ln(w_i(j_x)h + I^\star) + \dfrac{1}{2}\theta \right] & \frac{\theta - \underline{\varepsilon}}{\bar{\varepsilon} - \underline{\varepsilon}} > 1 \end{cases}$$

Choosing to stay single in period 1 means the agent will also be single in period 2. If a co-residential relationship was chosen in period 1, that relationship remains intact only if match quality is non-negative in period 2 after the match quality shock hits. Otherwise, the agent becomes single.[19] Those who cohabit in period 1 become married in period 2 if separation does not occur. Both marriage and cohabitation require an effort cost, with marriage requiring p[20] times more effort.

[19] This model has no explicit separation cost, although the implicit separation cost is the loss of a partner's income and match quality. While an explicit separation cost can be easily added into the model, the complication of the analytical solution is undesired.

[20] p in the data is the additional dates that those who marry without cohabiting go on, compared to those who cohabit.

From the agent's problem, $f(\varepsilon|\varepsilon \leqslant \theta) = (\theta - \underline{\varepsilon})/(\bar{\varepsilon} - \underline{\varepsilon})$ is interpreted as the probability that the shock received is less than match quality from period 1. If $(\theta - \underline{\varepsilon})/(\bar{\varepsilon} - \underline{\varepsilon}) < 0$, then even the smallest shock possible to match quality would cause the agent to have a negative match quality in period 2. Thus, the agent will always become single in period 2. On the other hand, if $(\theta - \underline{\varepsilon})/(\bar{\varepsilon} - \underline{\varepsilon}) > 1$, then even the biggest possible shock is not able to cause $\theta - \varepsilon < 0$, which means that the agent will always stay in the relationship regardless of the magnitude of the shock. For $0 \leqslant (\theta - \underline{\varepsilon})/(\bar{\varepsilon} - \underline{\varepsilon}) \leqslant 1$, the shock has a probability to be small enough that $\theta - \varepsilon > 0$, allowing for the relationship to continue.

The model assumes that if the agent chooses to cohabit or marry, that the partner will accept. This simplification greatly reduces the complexity of the analytical solution. I also suggest that the match quality each individual draws can account for the willingness of a potential partner to enter into a co-residential relationship. Match quality can be thought of as taking into account the probability that the partner will accept an offer to begin a co-residential relationship. Drawing a low match quality, then, can also be interpreted as being matched with someone who would not be willing to be in a co-residential relationship.

1.5 Analytical Solution

To solve for an analytical solution, consider the case $\theta \sim U\left[\underline{\theta}, \bar{\theta}\right]$, for $\underline{\theta} < 0$ and $\bar{\theta} > 1$, and $\varepsilon \sim U[0, 1]$. Then the agent's problem can be simplified to the following

$$\text{v(single)} = \ln(w_i(j_x)h) + \beta \ln(w_i(j_x)h)$$

$$\text{v(cohab)} = \begin{cases} -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \ln(w_i(j_x)h) & \theta < 0 \\ -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta & \\ \quad + \beta \left[\theta \left(\ln(w_i(j_x)h + I^\star) - p j_x h^\alpha + \frac{1}{2}\theta \right) \right. & 0 \leqslant \theta \leqslant 1 \\ \quad \left. + (1-\theta) \ln(w_i(j_x)h) \right] & \\ -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta & \\ \quad + \beta \left[\ln(w_i(j_x)h + I^\star) + \frac{1}{2}\theta - p j_x h^\alpha \right] & \theta > 1 \end{cases}$$

$$\text{v(marry)} = \begin{cases} -p j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \ln(w_i(j_x)h) & \theta < 0 \\ -p j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta & \\ \quad + \beta \left[\theta \left(\ln(w_i(j_x)h + I^\star) + \frac{1}{2}\theta \right) \right. & 0 \leqslant \theta \leqslant 1 \\ \quad \left. + (1-\theta) \ln(w_i(j_x)h) \right] & \\ -p j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta & \\ \quad + \beta \left[\ln(w_i(j_x)h + I^\star) + \frac{1}{2}\theta \right] & \theta > 1 \end{cases}$$

1.5.1 College graduates are less likely to cohabit

To show that college graduates are less likely to cohabit, I show that those with higher wages (college graduates) have a higher threshold match quality for cohabitation, but the threshold for marriage does not depend on income. Thus, the range of θ where cohabitation is optimal is smaller for college graduates.

An outline of the proof is as follows: I will first show that when $\theta < 0$, the agent always cohabits (Lemma 1) and when $\theta > 1$, the agent always marries (Lemma 2). Then, I show in Proposition 1 that the range of θ where cohabitation is optimal is smaller for those with higher wages, but the range that marriage is optimal does not depend on wages. The result that college graduates (those with higher wages) will cohabit less follows.

Lemma 1. *For $\theta < 0$, the value of marriage is always lower than the value of cohabitation.*

Proof.

$$\begin{aligned} & v(marry|\theta<0) - v(cohab|\theta<0) \\ &= -pj_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \ln(w_i(j_x)h) \\ &\quad - (-j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \ln(w_i(j_x)h)) \\ &= (1-p) j_x h^\alpha \end{aligned}$$

Since $1 < p < \frac{1}{1-\beta}$, $v(marry|\theta<0) - v(cohab|\theta<0) < 0$. □

Lemma 2. *For $\theta > 1$, the value of marriage is always higher than the value of cohabitation.*

Proof.

$$\begin{aligned} & v(\text{marry}|\theta>1) - v(\text{cohab}|\theta>1) \\ &= -pj_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \left[\ln(w_i(j_x)h + I^\star) + \frac{1}{2}\theta\right] \\ &\quad - \left(-j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \left[\ln(w_i(j_x)h + I^\star) + \frac{1}{2}\theta - pj_x h^\alpha\right]\right) \\ &= j_x h^\alpha \left(1 - p(1-\beta)\right) \end{aligned}$$

$$\begin{aligned} & \Rightarrow j_x h^\alpha \left(1 - \frac{1}{1-\beta}(1-\beta)\right) < j_x h^\alpha \left(1 - p(1-\beta)\right) < j_x h^\alpha \left(1 - (1-\beta)\right) \\ & \Rightarrow 0 < j_x h^\alpha \left(1 + p(-1+\beta)\right) < j_x h^\alpha \left(\beta\right) \\ & \Rightarrow 0 < v(\text{marry}|\theta>1) - v(\text{cohab}|\theta>1) < j_x h^\alpha \left(\beta\right) \end{aligned}$$

The third to last line is because $1 < p < 1/(1-\beta)$.

Thus, if $\theta > 1$, both college and non-college graduates will choose marriage over cohabitation. □

Proposition 1. *The range of θ where cohabitation is optimal is smaller for those with higher wages but the range of where marriage is optimal does not depend on wages.*

Proof. By Lemma 1, when $\theta < 0$, the value of cohabitation is always higher than the value of marriage, and by Lemma 2, when $\theta > 1$, the value of marriage is always higher than the value of cohabitation. Since the value of cohabiting and the value of marriage are continuous (proof found in appendix), and because the value of marriage is increasing at a greater rate in θ than the value of cohabitation (proof found in appendix), there must exist a $\theta^{\star}_{cm}$ such that the value of marriage and cohabitation are the same.

$$
\begin{aligned}
&-pj_x h^{\alpha} + \ln(w_i(j_x)h + I^{\star}) + \theta^{\star}_{cm} \\
&\quad + \beta \left[\theta^{\star}_{cm} \left(\ln(w_i(j_x)h + I^{\star}) + \frac{1}{2}\theta^{\star}_{cm} \right) + (1 - \theta^{\star}_{cm}) \ln(w_i(j_x)h) \right] \\
&= -j_x h^{\alpha} + \ln(w_i(j_x)h + I^{\star}) + \theta^{\star}_{cm} \\
&\quad + \beta \left[\theta^{\star}_{cm} \left(\ln(w_i(j_x)h + I^{\star}) - pj_x h^{\alpha} + \frac{1}{2}\theta^{\star}_{cm} \right) + (1 - \theta^{\star}_{cm}) \ln(w_i(j_x)h) \right] \\
&\Rightarrow -j_x h^{\alpha}(-1 + p - p\beta\theta^{\star}_{cm}) = 0 \\
&\Rightarrow \theta^{\star}_{cm} = \frac{p-1}{p\beta}
\end{aligned}
$$

$\theta^{\star}_{cm}$ does not depend on wage, and thus the threshold will be the same for both college and non-college graduates.

Given some $\widehat{\theta}$, it is optimal to cohabit if $\theta^{\star}_{sc}(w_i(j_x)) < \hat{\theta} < \theta^{\star}_{cm}$[21], where $\theta^{\star}_{sc}(w_i(j_x))$ is the threshold after which cohabitation is preferred to being single. I will show that $\theta^{\star}_{sc}(w_i(j_x))$ is increasing with wage, so that $\theta^{\star}_{sc}(w_{ncol}(j_x)) < \theta^{\star}_{sc}(w_{col}(j_x)) < \hat{\theta} < \theta^{\star}_{cm}$.

When $v(\text{single}) < v(\text{cohab}|\theta = 0)$, $\theta^{\star}_{sc}(w_i(j_x))$ is negative since $v(\text{cohab})$ is increasing with θ and $v(\text{single})$ does not change with θ. When $v(\text{cohab}|\theta = 0) \leqslant v(\text{single}) < v\left(\text{cohab}|\theta = (p-1)/p\beta\right)$, $\theta^{\star}_{sc}(w_i(j_x))$ is positive. Thus, $\theta^{\star}_{sc}(w_i(j_x))$ is defined as follows:

[21] I am only interested in the case where $\theta^{\star}_{sc}(w_i(j_x)) < \theta^{\star}_{cm}$ because only then would cohabitation ever be chosen. If $\theta^{\star}_{sc}(w_i(j_x)) \geqslant \theta^{\star}_{cm}$, then no cohabitation will happen, since after $\theta^{\star}_{cm}$ marriage is always preferred to cohabitation. Thus, it is assumed in this paper that the parameters are such that $\theta^{\star}_{sc}(w_i(j_x)) < \theta^{\star}_{cm}$.

$$\theta^{\star}_{sc}(w_i(j_x)) = \begin{cases} j_x h^{\alpha} - \left[\ln\left(\frac{w_i(j_x)h+I^{\star}}{w_i(j_x)h}\right)\right] & \text{if } v(\text{single}) < v(\text{cohab}|\theta=0) \\ \frac{-1}{\beta} + pj_x h^{\alpha} - \left[\ln\left(\frac{w_i(j_x)h+I^{\star}}{w_i(j_x)h}\right)\right] & \\ +\frac{1}{\beta}\sqrt{2\beta\left(j_x h^{\alpha} - \ln\left(\frac{w_i(j_x)h+I^{\star}}{w_i(j_x)h}\right)\right)} & \text{if } v(\text{cohab}|\theta=0) \leqslant v(\text{single}) \\ \overline{+\left(-1+\beta pjh^{\alpha} - \beta\ln\left(\frac{w_i(j_x)h+I^{\star}}{w_i(j_x)h}\right)\right)^2} & \leqslant v\left(\text{cohab}|\theta=\frac{p-1}{p\beta}\right) \end{cases}$$

An observation of $\theta^{\star}_{sc}(w_i(j_x))$ is that the term $j_x h^{\alpha} - \ln\left(\left(w_i(j_x)h + I^{\star}\right)/w_i(j_x)h\right)$ repeatedly shows up. This term is the difference between effort cost and income gain, which is the trade-off faced for those deciding between cohabitation and remaining single. As effort increases, so will $\theta^{\star}_{sc}(w_i(j_x))$, meaning that a greater θ is required before cohabitation is preferred to being single.

The term $\ln\left(\left(w_i(j_x)h + I^{\star}\right)/w_i(j_x)h\right)$ is the utility gain from having additional income from a co-residential relationship. The concavity of natural log implies for $w_{col}(j_x) > w_{ncol}(j_x)$:

$$\ln\left(\left(w_{col}(j_x)h + I^{\star}\right)/w_{col}(j_x)h\right) < \ln\left(\left(w_{ncol}(j_x)h + I^{\star}\right)/w_{ncol}(j_x)h\right)$$

Therefore, $j_x h^{\alpha} - \ln\left(\left(w_i(j_x)h + I^{\star}\right)/w_i(j_x)h\right)$ is increasing with wage. The higher the wage, the lower the gains from economies of scale, and the lower $\theta^{\star}_{sc}(w_i(j_x))$ will be.

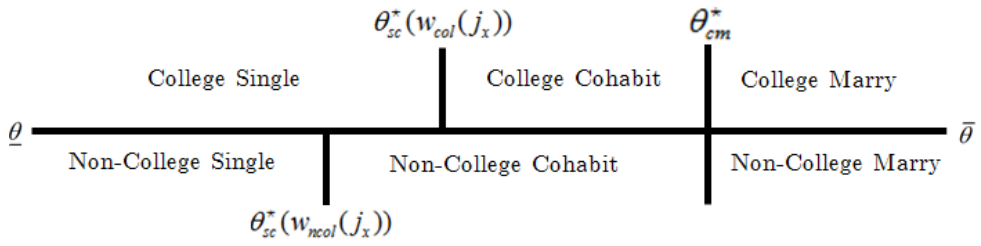
Thus, $\theta^{\star}_{sc}(w_{ncol}(j_x)) < \theta^{\star}_{sc}(w_{col}(j_x)) < \theta^{\star}_{cm}$.

□

Corollary 1. *College graduates (those with higher wages) cohabit less than non-college graduates.*

Proof. Proposition 1 shows that college graduates have a higher $\theta^{\star}_{sc}(w_i(j_x))$ than non-college graduates. It follows that since θ is drawn from the same distribution for both college and non-college graduates, that there is a lower probability that a college graduate will cohabit. Figure 1.2 summarizes the result.

Figure 1.2: College Graduates Cohabit Less Than Non-College Graduates



As a result, a lower percentage of college graduates will cohabit in period 1. □

1.5.2 College graduates are more likely to marry without cohabiting first

Corollary 2. *Marriage is a larger share of co-residential relationships for those with higher income.*

Proof. The result directly follows from Corollary 1. The percentage of co-residential relationships that are marriages in period 1 is total mar_{i}/ (total cohab_{i} + total mar_{i}) , and from Corollary 1, total $\text{cohab}_{\text{col}}$ < total $\text{cohab}_{\text{ncol}}$. Thus,

$$\%\ \text{mar}_{\text{col}} = \frac{\text{total mar}_{\text{col}}}{\text{total cohab}_{\text{col}} + \text{total mar}_{\text{col}}} > \frac{\text{total mar}_{\text{ncol}}}{\text{total cohab}_{\text{ncol}} + \text{total mar}_{\text{ncol}}} = \%\ \text{mar}_{\text{ncol}}$$

Thus, in period 1, a greater percentage of college graduates will choose to marry without cohabiting first.

□

1.5.3 College graduates are more likely to marry after cohabitation

Corollary 3. *More co-residential relationships transition into marriage for college graduates in period 2.*

Proof. If $\theta < 0$ upon entering period 2, since $\varepsilon \sim U[0,1]$, $\theta - \varepsilon < 0$ and the cohabitation will end with probability 1. This may happen as $\theta^{\star}_{sc}(w_i(j_x)) < 0$ in period 1. If $\theta > 0$, then there is a probability that the relationship will survive the shock. In period 2, the

probability of transitioning into marriage conditional on $\theta^{\star}_{sc}(w_i(j_x)) \leqslant \theta \leqslant \theta^{\star}_{cm}$ (having cohabited in period 1) is

$$P(\theta- > 0|\theta^{\star}_{sc}(w_i(j_x)) \leqslant \theta \leqslant \theta^{\star}_{cm}) = \int_{\max\{0,\theta^{\star}_{sc}(w_i(j_x))\}}^{\theta^{\star}_{cm}} \int_{0}^{=\theta} \frac{1}{\theta^{\star}_{cm} - \theta^{\star}_{sc}(w_i(j_x))} dd\theta$$

$$= \begin{cases} \frac{1}{2} \frac{{\theta^{\star}_{cm}}^2}{\theta^{\star}_{cm} - \theta^{\star}_{sc}(w_i(j_x))} & \text{if } \theta^{\star}_{sc}(w_i(j_x)) < 0 \\ \frac{1}{2} \left(\theta^{\star}_{cm} + \theta^{\star}_{sc}(w_i(j_x))\right) & \text{if } \theta^{\star}_{sc}(w_i(j_x)) \geqslant 0 \end{cases}$$

There are three cases to consider:

- Case 1: $\theta^{\star}_{sc}(w_{ncol}(j_x)) < \theta^{\star}_{sc}(w_{col}(j_x)) < 0$

  If cohabitation threshold for both college and non-college graduates are below 0, then

$$P(\theta- > 0|\theta^{\star}_{sc}(w_i(j_x)) \leqslant \theta \leqslant \theta^{\star}_{cm}) = \frac{1}{2} \frac{{\theta^{\star}_{cm}}^2}{\theta^{\star}_{cm} - \theta^{\star}_{sc}(w_{col}(j_x))}$$

$$> \frac{1}{2} \frac{{\theta^{\star}_{cm}}^2}{\theta^{\star}_{cm} - \theta^{\star}_{sc}(w_{ncol}(j_x))} = P(\theta- > 0|\theta^{\star}_{sc}(w_i(j_x)) \leqslant \theta \leqslant \theta^{\star}_{cm})$$

  Since $\theta^{\star}_{sc}(w_{ncol}(j_x)) < \theta^{\star}_{sc}(w_{col}(j_x)) < 0$, $\theta^{\star}_{cm} - \theta^{\star}_{sc}(w_{col}(j_x)) < \theta^{\star}_{cm} - \theta^{\star}_{sc}(w_{ncol}(j_x))$, hence the result above. Since the threshold for cohabitation is higher for college graduates, a higher proportion of cohabitations will transition into marriage for college graduates.

- Case 2: $\theta^{\star}_{sc}(w_{ncol}(j_x)) < 0 \leqslant \theta^{\star}_{sc}(w_{col}(j_x))$

  If the cohabitation threshold is greater than or equal to zero for college graduates but negative for non-college graduates, then the probability of transition into marriage is higher for college graduates:

$$P(\theta - \varepsilon > 0|\theta^{\star}_{sc}(w_i(j_x)) \leqslant \theta \leqslant \theta^{\star}_{cm}) = \frac{1}{2} \left(\theta^{\star}_{cm} + \theta^{\star}_{sc}(w_{col}(j_x))\right)$$

$$> \frac{1}{2} \frac{{\theta^{\star}_{cm}}^2}{\theta^{\star}_{cm} - \theta^{\star}_{sc}(w_{ncol}(j_x))} = P(\theta - \varepsilon > 0|\theta^{\star}_{sc}(w_i(j_x)) \leqslant \theta \leqslant \theta^{\star}_{cm})$$

To see the result, note that $\theta^{\star}_{sc}(w_{ncol}(j_x))$ has an upper bound of zero, which means $(1/2)\theta^{\star\,2}_{cm}/\theta^{\star}_{cm} - \theta^{\star}_{sc}(w_{ncol}(j_x))$ has an upper bound of $(1/2)\theta^{\star\,2}_{cm}/\theta^{\star}_{cm} = (1/2)\theta^{\star}_{cm}$. Since $(1/2)\left(\theta^{\star}_{cm} + \theta^{\star}_{sc}(w_{col}(j_x))\right) > (1/2)\theta^{\star}_{cm}$, the above inequality holds.

- Case 3: $0 \leqslant \theta^{\star}_{sc}(w_{ncol}(j_x)) \leqslant \theta^{\star}_{sc}(w_{col}(j_x))$

  Proposition 1 showed that $\theta^{\star}_{sc}(w_{col}(j_x)) > \theta^{\star}_{sc}(w_{ncol}(j_x))$, therefore

$$P(\theta - \varepsilon > 0 | \theta^{\star}_{sc}(w_i(j_x)) \leqslant \theta \leqslant \theta^{\star}_{cm}) = \frac{1}{2}\left(\theta^{\star}_{cm} + \theta^{\star}_{sc}(w_{col}(j_x))\right)$$
$$> \frac{1}{2}\left(\theta^{\star}_{cm} + \theta^{\star}_{sc}(w_{ncol}(j_x))\right) = P(\theta - \varepsilon > 0 | \theta^{\star}_{sc}(w_i(j_x)) \leqslant \theta \leqslant \theta^{\star}_{cm})$$

  Thus, college graduates are more likely to transition into marriage from cohabitation for $0 < \theta^{\star}_{sc}(w_{col}(j_x))$.

□

1.5.4 Working more increases transition probability into marriage

Holding all else equal, those who work more prior to cohabitation will have a higher threshold for cohabitation, since working more increases effort to begin a relationship. However, the threshold for marriage does not depend on hours worked, as shown in Proposition 1. Thus, I show that agents who work more have a higher $\theta^{\star}_{sc}$, and it follows from Corollary 3 that those with a higher $\theta^{\star}_{sc}$ in period 1 will have a higher probability of marriage in period 2.

Proposition 2. *Suppose* $\bar{h} > \underline{h} > 0$ *, then* $\theta^{\star}_{sc,\bar{h}}(w_i(j_x)) > \theta^{\star}_{sc,\underline{h}}(w_i(j_x))$.

Proof. In the definition of $\theta^{\star}_{sc}(w_i(j_x))$, the terms that include the expression $j_x h^{\alpha} - [\ln(w_i(j_x)h + I^{\star})/\ln(w_i(j_x)h)]$[22] are the only terms that have hours worked h. Therefore, finding the sign of $\delta j_x h^{\alpha} - [\ln(w_i(j_x)h + I^{\star})/\ln(w_i(j_x)h)]/\delta h$ will determine whether $\theta^{\star}_{sc}(w_i(j_x))$ is increasing or decreasing in hours worked.

[22] Note that in $\theta^{\star}_{sc}(w_i(j_x))$, $pj_x h^{\alpha} - [\ln(w_i(j_x)h + I^{\star})/\ln(w_i(j_x)h)]$ also appears, but that will also increase with hours worked:

$$\frac{\delta p j_x h^{\alpha} - \left[\frac{\ln(w_i(j_x)h + I^{\star})}{\ln(w_i(j_x)h)}\right]}{\delta h} = \frac{w_i(j_x)}{w_i(j_x)h} - \frac{w_i(j_x)}{w_i(j_x)h + I^{\star}} + \alpha p j_x h^{-1+\alpha} > 0$$

It can be shown that the expression is increasing with hours worked:

$$\frac{\delta j_x h^{\alpha} - \left[\frac{\ln(w_i(j_x)h+I^{\star})}{\ln(w_i(j_x)h)}\right]}{\delta h}$$
$$= \frac{w_i(j_x)}{w_i(j_x)h} - \frac{w_i(j_x)}{w_i(j_x)h + I^{\star}} + \alpha j_x h^{-1+\alpha} > 0$$

Since the expression is positive and always added in the definition of $\theta^{\star}_{sc}(w_i(j_x))$, it implies that $\delta\theta^{\star}_{sc}(w_i(j_x))/\delta h > 0$. Thus,

$$\theta^{\star}_{sc,\bar{h}}(w_i(j_x)) > \theta^{\star}_{sc,h}(w_i(j_x))$$

$\theta^{\star}_{cm}$ does not depend on h , so it follows that $\theta^{\star}_{sc,\bar{h}}(w_i(j_x)) < \theta^{\star}_{sc,h}(w_i(j_x)) < \theta^{\star}_{cm}$ because of the earlier assumption that $\theta^{\star}_{sc}(w_i(j_x)) < \theta^{\star}_{cm}$. In period 2, because average match quality will be higher for those who work $\bar{h}$, they will have a higher probability of transitioning into marriage. □

1.5.5 Agents with non-traditional job schedules are more likely to only cohabit

Since $w_i(j_f) > w_i(j_{uf})$ for $i = \{col, ncol\}$, having a non-traditional work schedule, j_{uf}, will cause agents to gain more from economies of scale in a relationship. It follows from Corollary 1 that those with low wages will cohabit more than those with high wages. However, having a non-traditional work schedule will also increase effort, since $e(h, j_{uf}) > e(h, j_f)$. In the data, the percentage decrease in wage from having a non-traditional work schedule is greater than the percentage decrease in the number of dates. Therefore, I assume that a non-traditional schedule will decrease wage more than it will increase effort.

$$j_{uf}h^{\alpha} - [\ln(w_i(j_{uf})h + I^{\star})/\ln(w_i(j_{uf})h)]$$
$$< j_f h^{\alpha} - [\ln(w_i(j_f)h + I^{\star})/\ln(w_i(j_f)h)]$$
$$\Rightarrow \theta^{\star}_{sc}(w_i(j_{uf})) < \theta^{\star}_{sc}(w_i(j_f))$$

Therefore, agents with non-traditional job schedules will cohabit more than those with a traditional job schedule.

1.6 Empirical Evidence

In this paper, the main mechanism by which college and non-college graduates differ in cohabitation behavior is through the differing value of economies of scale. One of the main ways in which a couple can take advantage of economies of scale is through sharing the cost of housing. In this section, I document evidence that county-level cohabitation rate relative to living alone increases with cost of housing.

I use the U.S. Census Bureau American Community Survey (ACS) accessed through the Minnesota Population Center's IPUM-USA harmonized database[23] for county-level cohabitation rate and household-level characteristics. Importantly, the ACS provides a household roster for each household. Through the roster I observe the composition of each household, without which it would be unclear how cohabitation rate changes with housing cost. As an example, a single individual living by himself may choose to live with roommates or cohabit as rent increases to save on the cost of housing. However, someone who is single and already living with roommates would not need to cohabit to take advantage of economies of scale. If economies of scale is the main reason why single individuals choose to cohabit, increasing housing cost should mainly affect individuals who are living alone and not those who are living with others already. Thus, I want to see the percentage of single individuals that go from living alone to cohabiting as housing cost increases, rather than just single individuals in general.

I classify those between ages 21 and 45 as either living alone or living with someone else in the household[24] (not living alone). For those who are single, living alone means they are the only member of their household, while not living alone means there is at least one other person in the same household. A cohabiting couple is living alone if the household roster only contains two individuals, the householder and the householder's cohabiting partner, while cohabitors who are not living alone are those with additional household members. Lastly, for married couples, living alone implies the householder and the householder's spouse are the only two members of the household, while married couples not living alone means the married couple is in a household of more than two

[23] Ruggles et al. (2015)

[24] To be consistent with the data presented earlier in this paper, only households without children are considered. Thus, "someone else in the household" implies there is a relative or roommate, friend, or boarder in the same household.

people. Table 1.11 summarizes this classification.

Table 1.11: Classification of Living Alone or Not Living Alone

	Living Alone	Not Living Alone
Single	Only one individual in household, and that individual is single	Individual is single and lives in a household with more than one member, but not own children
Cohabit	Only two individuals in household; householder and cohabiting partner	Householder and cohabiting partner live with other related or unrelated individuals, but not own children
Marry	Only two individuals in household; householder and married spouse	Householder and married spouse live with other related or unrelated individuals, but not own children

I use Zillow's ZRI time series for rental data for single family residences, condos, and co-ops as my primary measure of the cost of housing in each county. For comparison, I also use Zillow's ZHVI All Homes time series, which contains median sale prices for single family residences, condos, and co-ops and Zillow's ZRI per square foot time series. A useful feature of the Zillow data is that calculations of rent or median sale prices are based on a constant "basket" of homes, much like the consumer price index measures the change in the cost of a constant basket of goods. This is a more accurate measure of the changes in the cost of housing because the composition of housing available is held fixed.

The first complete year in the Zillow data is 2011 and the last is 2014, thus these four years were used from the ACS[25]. I focus particularly on four household types: single

[25] Zillow indexes are computed monthly. The monthly data was averaged for each year to get an annual figure.

and living alone, single and not living alone, cohabiting and living alone, and married and living alone.

In the ensuing analysis, the dependent variable is the proportion of the county of a specific education group that is single and living alone, single and not living along, cohabiting and living alone, or married and living alone. The explanatory variable is education interacted with a measure of housing cost. Regression (1) uses average monthly rent, regression (2) uses median housing sale price, and regression (3) uses average rent per square foot. Following Aitchison (1986)'s solution to compositional data problems, I use a log-ratio transformation with single and living alone as the base group. County-level fixed effects are also used. Tables 1.12, 1.13, and 1.14 show the results.

Table 1.12: Log Ratio: Single (Not Alone) to Single (Alone)

	(1)		(2)		(3)	
	NC	C	NC	C	NC	C
Average Monthly Rent	.0005***	.0002				
	(.0002)	(.0002)				
Median Sale Price			1.22e-06**	1.23e-06**		
			(5.36e-07)	(5.37e-07)		
Median Rent Per Square Feet					.87***	.55**
					(.28)	(.28)
Intercept	-.23***		.05		-.33***	
	(.17)		(.07)		(.17)	

*** denotes statistical significance at a 1% level, ** denotes significance at the 5% level, and * denotes significant at the 10% level. Controlled for county-level fixed effects.

Table 1.13: Log Ratio: Cohabiting (Alone) to Single (Alone)

	(1)		(2)		(3)	
	NC	C	NC	C	NC	C
Average Monthly Rent	.0008***	.0002				
	(.0003)	(.0003)				
Median Sale Price			1.72e-06**	1.17e-06		
			(8.42e-07)	(8.49e-07)		
Median Rent Per Square Feet					1.29***	.42
					(.42)	(.44)
Intercept	-2.12***		-1.75***		-2.25***	
	(.27)		(.11)		(.27)	

*** denotes statistical significance at a 1% level, ** denotes significance at the 5% level, and * denotes significant at the 10% level. Controlled for county-level fixed effects.

Table 1.14: Log Ratio: Married (Alone) to Single (Alone)

	(1)		(2)		(3)	
	NC	C	NC	C	NC	C
Average Monthly Rent	.0002	.0004*				
	(.0002)	(.0002)				
Median Sale Price			1.16e-06*	3.52e-08		
			(6.34e-07)	(6.34e-07)		
Median Rent Per Square Feet					.38	.54
					(.33)	(.33)
Intercept	-1.00***		-0.72***		-1.01***	
	(.21)		(.08)		(.20)	

*** denotes statistical significance at a 1% level, ** denotes significance at the 5% level, and * denotes significant at the 10% level. Controlled for county-level fixed effects.

A positive coefficient implies that as the measure of the cost of housing increases, so

does the log-ratio between the proportion of a specific household type and the proportion of singles living alone in that county. In this analysis, I focus on the direction of the movement rather than the exact level of change. Recall from the theory section above that if cost of housing increases, I expect there to be more cohabitors but not more married individuals.

In Table 1.12, as average monthly rent increases, so does the proportion of singles not living alone to the proportion of singles living alone. This effect for non-college graduates is statistically significant. In Table 1.13, changes in average rent has a significant effect on the ratio of cohabitors to singles living alone for non-college graduates. As my theory suggests, non-college graduates are those with lower income, and thus the ones who have a higher value of economies of scale.

Table 1.14 shows that there is not a significant effect of average rent on the married to single and living alone ratio for non-college graduates. One of the results in my theory is that economies of scale will not drive individuals to marriage, since cohabitation has the same economies of scale benefit as marriage and requires less effort to begin. Marriage will still only occur if match quality is high enough. Assuming that average rent and the distribution of match quality are independent, the proportion of individuals married should not change as average rent changes. While Table 1.14 shows that this is so for non-college graduates, average rent seems to have a positive and significant effect for college graduates marriage rates. This suggests that my theory may be more suited for analyzing individuals with low income, such as non-college graduates.

Using median sale price of houses or average rent per square foot yield results in the same direction as using average rent. Importantly, all three measures suggests that the proportion cohabiting tends to increase relative to the proportion of singles living alone as cost of housing increases.

1.7 Discussion

1.7.1 Increase in cohabitation and decrease in marriage

The rise of cohabitation rates and fall of marriage rates in recent decades have been well-documented, such as in Bumpass and Lu (2000), Fields and Casper (2001), and Stevenson and Wolfers (2007). Stevenson and Wolfers suggest that possible driving

forces include diminishing social stigma and changing of policies[26] that decreased the value of marriage. As an application of my theory, I propose that an additional factor is the declining of real wages.

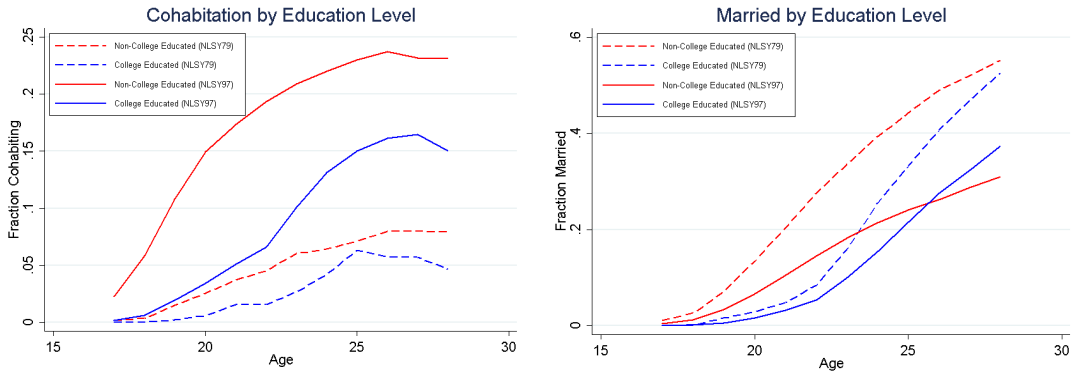
The left panel on Figure 1.3 shows that cohabitation rate (proportion currently cohabiting) at each age had increased for both college graduates and non-college graduates between the two NLSY cohorts (1979 and 1997), and the right panel shows that marriage rate has fallen for both. Particularly, the increase in cohabitation rate and the decrease in marriage rate had been greater for non-college graduates than for college graduates. The NLSY79 cohort was born between the years 1957-1964 (compared to 1980-1984 for the NLSY97 cohort). A comparison in cohabitation and marriage rates between the NLSY79 and NLSY97 cohort, then, is an estimate of the change in cohabitation and marriage rates in twenty years.

According to my mechanism, a decrease in real wages for both college and non-college graduates will lead to more cohabitation and less marriage for both groups. Additionally, if the decrease in real wages for non-college graduate is greater than that of college graduates, then the increase in cohabitation and decrease in marriage will be more drastic for non-college graduates. Kambourov and Manovskii (2009) documents exactly this in their paper. Real wages over the life cycle for both college and non-college groups had been flattening for five cohorts between the years 1968-1988, but the non-college group had a significantly greater flattening of wages.

This suggests that the increase in cohabitation and decrease in marriage is not exclusively due to a change in culture, social norms, and policies over the last five decades. The changing labor market characteristics of individuals over the last five decades has also changed the incentives of relationship formation.

[26] Such as Choo and Siow (2006) which suggests that legalization of abortion decreased the gains from marriage, contributing to its decline.

Figure 1.3: Cohabitation and Marriage Rate for NLSY79 and NLSY97 Cohorts, by Education



(a) Percent Cohabiting

(b) Percent Married

1.7.2 Welfare Implications

As an application of my theory, I qualitatively evaluate how programs targeting urban poverty will impact relationship formation. My mechanism suggests that cohabitation rate among the poor is high because they have low wages, which having low education or being less likely to have traditional work schedules contributes to. Additionally, even though I assumed in my model that college and non-college graduates do not differ in their hours worked, it can be seen in the data that college graduates work more than non-college graduates[27]. These factors lead low income individuals to be less selective in choosing their potential partners. Therefore, programs that increase wages or hours worked for individuals are predicted to decrease cohabitation rates. Importantly, my mechanism suggests that the decrease in cohabitation rate as a result of higher wages or hours worked is completely from the decline in formation of relationships with low match quality.

This decline has implication for welfare. For instance, Johnson and Ferraro (2000) and Gaertner and Foshee (1999) suggest that there is a negative correlation between

[27] Adding a difference in hours worked for college and non-college graduates will not change my results at all, since my results are qualitative. Adding the difference in hours worked will make non-college graduates even less selective over match quality and college graduates to be even more selective over match quality.

commitment in a relationship and domestic violence. In my mechanism, it is predicted that cohabiting relationships with low match quality will have a higher probability of dissolution. On the other hand, couples with high match quality have a higher probability of transitioning into marriage. Thus, match quality is correlated with relationship commitment, and a decline in low-match cohabitation would occur with a decline in relationships with low commitment, leading to a decrease in domestic violence.

Welfare is also affected by changes to fertility and investment in children. An increasing percentage of non-marital births are born to couples who are cohabiting. Lichter (2012) documents that almost 60% of non-marital births in 2006-2008 were born to cohabiting couples. Martinez (2015) documents that while the percentage of fathers whose first child is born out of wedlock has decreased between 1980 and 2009, the percentage of fathers whose first child is born within a cohabiting relationship had increased. There is also a higher likelihood that a child born within cohabitation was a result of an unintended pregnancy. Finer and Henshaw (2006) finds that the unintended birth rate of cohabiting women is more than twice of married women or of unmarried women who are not cohabiting. Additionally, the unintended birth rate is higher among the poor. Since there is a higher probability that low-income cohabitors will have their relationships end in dissolution[28], low-match cohabitations among the poor are more likely to result in formation of single-parent households. A decline in low-match cohabitation therefore should lead to a decline in the number of single parent households.

While the costs and benefits of urban poverty programs need to be considered quantitatively to make policy recommendations, my results highlight that the evaluation of such policies should not ignore the changes in welfare due to changed incentives of relationship formation.

1.8 Conclusion

In this paper, I study formation of cohabitation and marriage and focus on how individual labor market decisions and characteristics prior to the start of cohabitation affect

[28] Lichter et al. (2006) finds that it is more likely that cohabitation end within five years by dissolution rather than marriage, especially for the low-income. Of low-income cohabiting relationships that end, 63% end by dissolution (37% transition into marriage).

the transition of cohabitation into marriage. I find that an individual's number of hours worked and wages prior to cohabitation influence this transition probability. Lower wages are associated with lower probability of transitioning into marriage, whereas high number of hours worked is associated with a higher transition probability. I propose a theory of relationship formation consistent with the data, and provide a two-period model to give intuition.

Cohabitation and marriage both provide economies of scale and relational utility, but there is a disutility from effort to begin either. This is the basic trade-off that individuals face when deciding to form a co-residential relationship. Since the gains in economies of scale is the same for cohabitation and marriage and the effort needed to be exerted prior to marriage is higher, marriage will only be optimal if match quality is sufficiently high. Those with low income select into cohabiting relationships instead of remaining single due to the high value of economies of scale. Those with high disutility from exerting pre-relationship effort, such as individuals who have high number of hours worked, are more selective over match quality for two reasons: 1) because it is more dis-utilizing to begin a relationship, and 2) because their value for economies of scale is lower due to a higher income. This implies that if they cohabit, they will have higher match quality and thus are more likely to transition into marriage.

I document on a county-level that as the cost of housing increases, so does the relative proportion of non-college graduates cohabiting to those single and living alone. While the relative proportion of cohabitors to single individuals living alone move in the same direction for college-graduates, the results were not statistically significant. This suggests that my theory may be more useful in analyzing low income individuals.

I suggest that decreasing real wages can contribute to the increasing cohabitation rate and decreasing marriage rate over the last five decades. I also suggest that changes to the incentives to form a co-residential relationship create an additional channel where urban poverty programs can affect welfare.

The theory presented in this paper does not take into account a few important aspects of co-residential relationships. The first is work or tax benefits that can affect incentives to cohabit or marry. For example, while virtually all employers will allow for health insurance to be extended to a spouse, that is not the case for a cohabiting partner. Or consider the example where by marrying, two individuals move into a higher

tax bracket and thus is taxed more because of marriage. This penalty would also affect the decision of individuals to marry or cohabit[29]. While these extensions can be added into the model, a careful analysis of tax codes and employer benefits is needed, and that is beyond the scope of this paper. A careful analysis of how tax and benefits change with marital status and its impact on relationship formation remains an important topic for future research.

[29] The impacts of tax policies on marriage and cohabitation had been addressed in the literature, see Chade and Ventura (2005) and Light and Omori (2008). Adding tax into the model that depends on marital status will give a similar result to Chade and Ventura (2005), where it increases the threshold match quality to marry if there is a "marriage penalty", and thus increases the number of cohabiting relationships. However, to my knowledge, no paper in the literature uses a monthly panel that can track precisely how incentives to marriage or cohabitation change.

Chapter 2

Moving In or Moving On?

2.1 Introduction

It has been well documented that cohabitation in the United States has been on the rise in the last four decades (e.g. Bumpass and Sweet (1995), Bumpass and Lu (2000), Fields and Casper (2001), and Stevenson and Wolfers (2007)). This has been the case for both non-college and college graduates, though non-college graduates still tend to cohabit more than college graduates. The change and difference in migration is explored in this paper as a possible contributor to both the change over time and the cross-sectional difference in cohabitation rate between the two education groups. The intuition is as follows: An individual who thinks they have a high probability of having a big move, such as one out of county or state, may be less willing to settle down by cohabiting. Thus, college graduates moving more than non-college graduates can lead to a lower cross-sectional cohabitation rate for college graduates. Additionally, a decline in migration rates for both education groups in the last three decades could then increase cohabitation rate over time.

To test this hypothesis, I compare cohabitation and migration behavior between the National Longitudinal Survey of Youth (NLSY) 1979 and 1997 cohorts. The NLSY97 cohort was born about 20 years after the NLSY79 cohort. I find that non-college graduates cohabit more than college graduates between the ages 21 and 28 in both cohorts. I also document that cohabitation rates are higher in the NLSY97 cohort for both education groups. I then show that the ratio of college graduate cohabitors to

non-college graduate cohabitors has become closer to one in the NLSY97 cohort. This suggests that college graduates are more similar to non-college graduates with regards to cohabitation in the NLSY97 cohort compared to the NLSY79 cohort.

I use the geo-coded NLSY data to show that migration rates are higher for college graduates than non-college graduates in both cohorts. I estimate the probability of moving and find that college graduates are more likely to have a big move than non-college graduates. I also show that migration rates are more similar between non-college and college graduates in the NLSY97 cohort than in the NLSY79 cohort.

I then estimate an individual's odds of cohabiting on whether the individual had a big move in the next year. I use two different measures to define what a big move looks like. First, it can be an out of county or state move, and second, it can be a move of greater than 50 miles. For both measures, I find that there is a decrease in the log odds of cohabiting in the current year if there is a big move coming up in the following year. This suggests that the difference in migration behavior could explain the cross-sectional difference in cohabitation rate between non-college and college graduates within each cohort. Additionally, I find that the median college graduation age of 23 is when the migration rate between college and non-college graduates is most dissimilar, while ages 21 and 22 are when cohabitation rate is the most dissimilar.

With regards to the changes in migration over time, the NLSY yielded a surprising result in that migration rates were higher in the 1997 cohort than the 1979 cohort. This contradicts with what is well documented in the literature (e.g. Molloy et al. (2011), and Kaplan and Schulhofer-Wohl (2015)). I use the Census Bureau's Current Population Survey to show that migration out of county and state for those between 21 and 28 indeed had declined between the two cohorts. This implies that migration cannot be ruled out as having contributed to the increase in cohabitation rates in the last four decades.

The paper proceeds as follows. Section 2 describes the data from the NLSY79 and NLSY97 and presents several findings on cohabitation and migration. Section 3 shows the result of future migration lowering the odds of cohabiting. Section 4 discusses the data and results in the context of this paper's hypothesis. Section 5 concludes.

2.2 Data

I use the National Longitudinal Survey of Youth 1979 (NLSY79) and the National Longitude Survey of Youth 1997 (NLSY97), which are two panel data sets following youths into adulthood. The NLSY79 was administered annually until 1994, when it switched to being conducted bi-annually. The NLSY97 is a monthly event history survey administered yearly. The NLSY79 cohort was born between 1957-1964, and comprises of 12,686 youths that were first surveyed in 1979. The NLSY97 cohort was born between 1980-1984, and comprises of 8984 youths that were first surveyed in 1997. I analyze individuals between the ages of 21 and 28[1].

The restricted geocoded data for both NLSY79 and NLSY97 was used to provide additional information on individual location and distance moved since the last interview year. Individual location is given at the county level, while distance moved is calculated either from exact distances between current address and previous address (when available) or the distance between the centroids of the new and old zip codes. While moving history is available in the publicly available NLSY97 data, it is not for the NLSY79. Therefore, I construct a yearly moving history for each individual with a method that is consistent for both the NLSY79 and NLSY97.

The method is as follows. For each individual in either the NLSY79 or NLSY97, I compare his or her current and previous interview year's county and state. If just the county is different but the state is the same, then it is considered an out of county move. If both county and state are different, then it is considered an out of state move. The variable of interest in my analysis is the combination of the two – moves that are out of county or out of state. If any location data is missing for a person-year, then the individual neither moved nor not moved. Summary statistics by cohort and education are shown in Table 2.1.

The numbers shown in the table are in person-years. Each individual in the NLSY79 or NLSY97 has a maximum of 8 person-years of data. As an example of how to interpret Table 2.1, there were 39,127 person-years where non-college graduates did not move

[1] The NLSY97 round 15 is the last round used in my data. The maximum age of 28 is used only because of the availability of data, and does not have any significance. There are less observations in the NLSY97 starting at age 28, and the lack of data from ages 29-32 can skew the results for those ages.

and 4,332 person-years where non-college graduates moved out of county or state in the NLSY79. A non-college graduate in the NLSY79 cohort who moved out of county or state at least once every interview year between ages 21 and 28 would account for 8 of 4,332 moves recorded, but zero of 39,127 non-moves recorded. A non-college graduate in the NLSY79 cohort who did not move at all in the 8 years between ages 21 and 28 would account for 0 of 4,332 moves recorded, but 8 of 39,127 non-moves recorded.

Table 2.1: Out of County or State Moves by Cohort and Education (Person-Years)

	Non-College Graduate	College Graduate
NLSY79		
No Out of County/State Move	39127	6312
Out of County/State Move	4332	1345
NLSY97		
No Out of County/State Move	35234	12038
Out of County/State Move	5291	2799

An observation from Table 2.1 is that college graduates have a higher probability of moving than non-college graduates in both the NLSY79 and NLSY97 cohorts. A comparison between the NLSY79 and NLSY97 will be made later in Section 2.2.2.

2.2.1 Difference in Cohabitation Rate

Figure 2.1 shows the cohabitation rate for the NLSY79 and NLSY97 cohorts by education and between the ages of 21 and 28. Two main observations can be made from Figure 2.1. First, college graduates are less likely to cohabit between the ages of 21 and 28 in both the NLSY79 and NLSY97 cohorts. Second, cohabitation rate has increased between the NLSY79 and NLSY97 cohorts for both college and non-college graduates.

Figure 2.1: Cohabitation Rate by Education

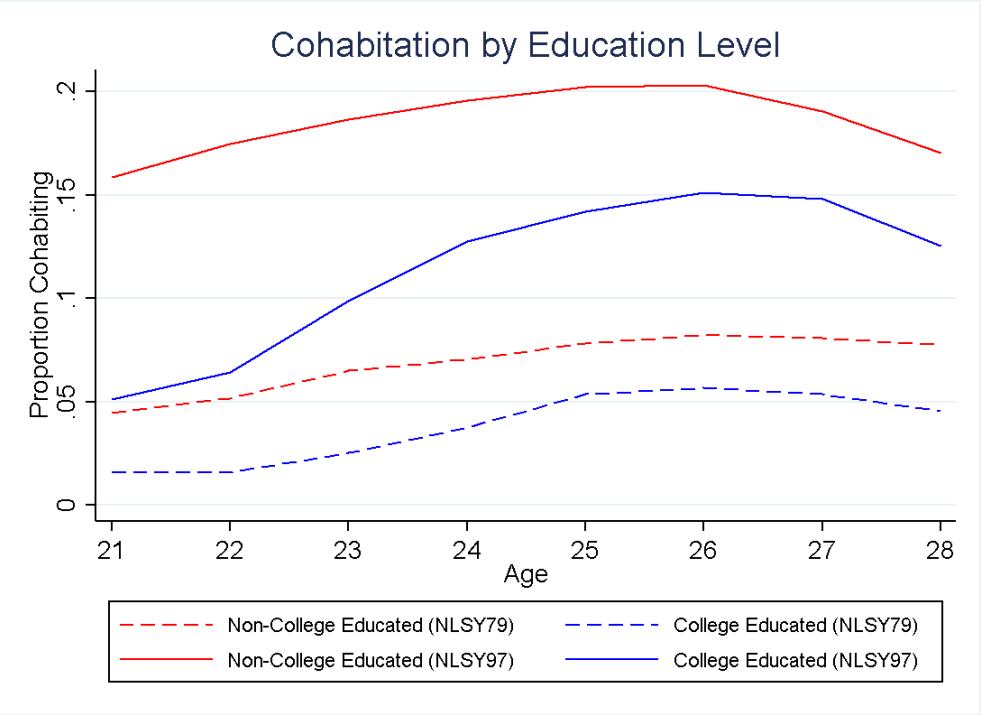
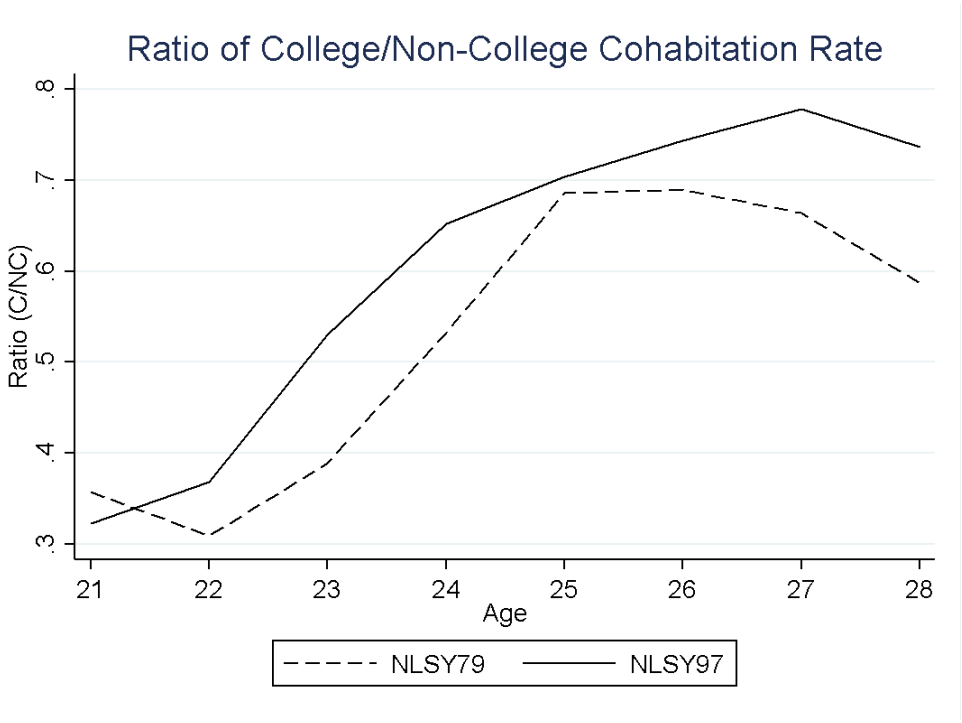


Figure 2.2 shows the ratio of college graduate to non-college graduate cohabitation rate. The ratio is defined as college graduate cohabitation rate divided by non-college graduate cohabitation rate. Thus, the closer to one the ratio is, the more similar non-college graduates and college-graduates are in cohabitation rate. It can be seen that from the NLSY79 cohort to the NLSY97 cohort, the relative cohabitation ratio has become closer to one. This means that the difference between college and non-college graduate cohabitation rate has fallen. Additionally, the age at which non-college graduates are cohabiting most relative to college graduates is 22 for the NLSY79 and 21 for the NLSY97.

Figure 2.2: Ratio of College to Non-college Cohabitation Rate



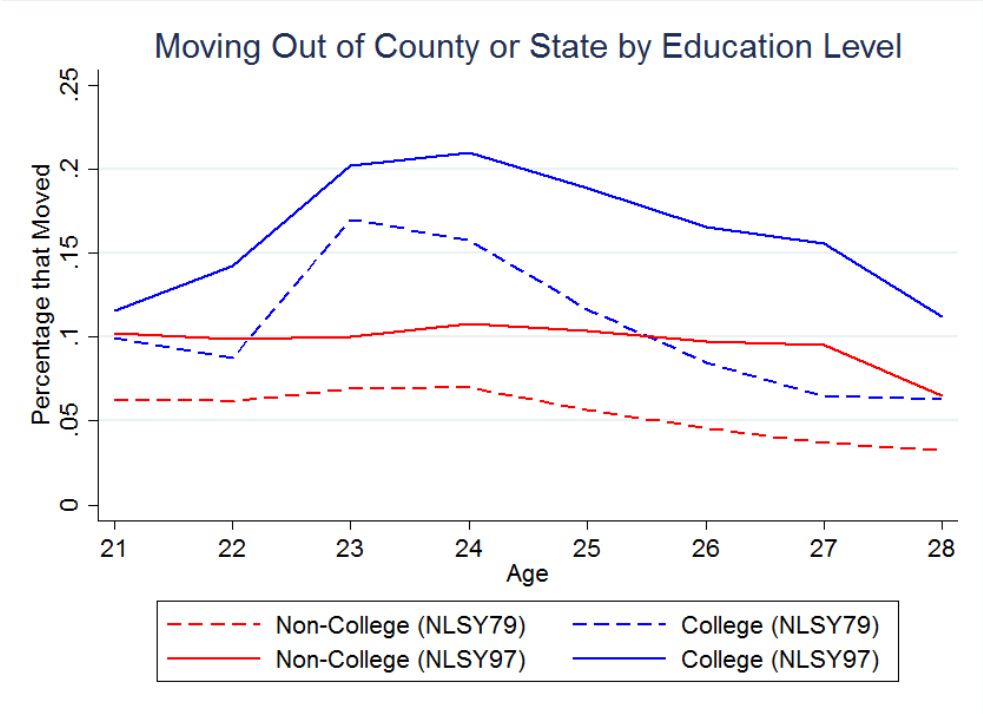
2.2.2 Difference in Migration Rate

Figure 2.3 shows the percentage of non-college graduates and college graduates in either the NLSY79 or NLSY97 who moved out of county or state at least once at each age between 21 and 28. College graduates in both cohorts have a higher probability of moving between the ages of 21 and 28 when compared to non-college graduates. Comparing between cohorts, college graduates in the NLSY97 cohort tend to have a higher probability of moving than their counterparts in the NLSY79 cohort. This is also true of non-college graduates in the NLSY97 cohort. Thus, the NLSY data suggests that out of county or state migration is higher for the NLSY97 cohort than the NLSY79 cohort.

The results of a probit regression where the probability of moving out of county or state for the NLSY79 cohort is regressed on whether the individual is a college graduate, age, number of children in household, and real hourly wage is shown in the left panel of Table 2.2, under regression (1). Regression (2) uses a different measure of migration – whether or not a far move occurred. A far move is defined here as a move that is

greater than 50 miles. The right panel of Table 2.2 shows the results of the same two regressions for the NLSY97 cohort. For both cohorts, the results confirm that being a college graduate increases the likelihood of having a big move.

Figure 2.3: Moving Out of County or State



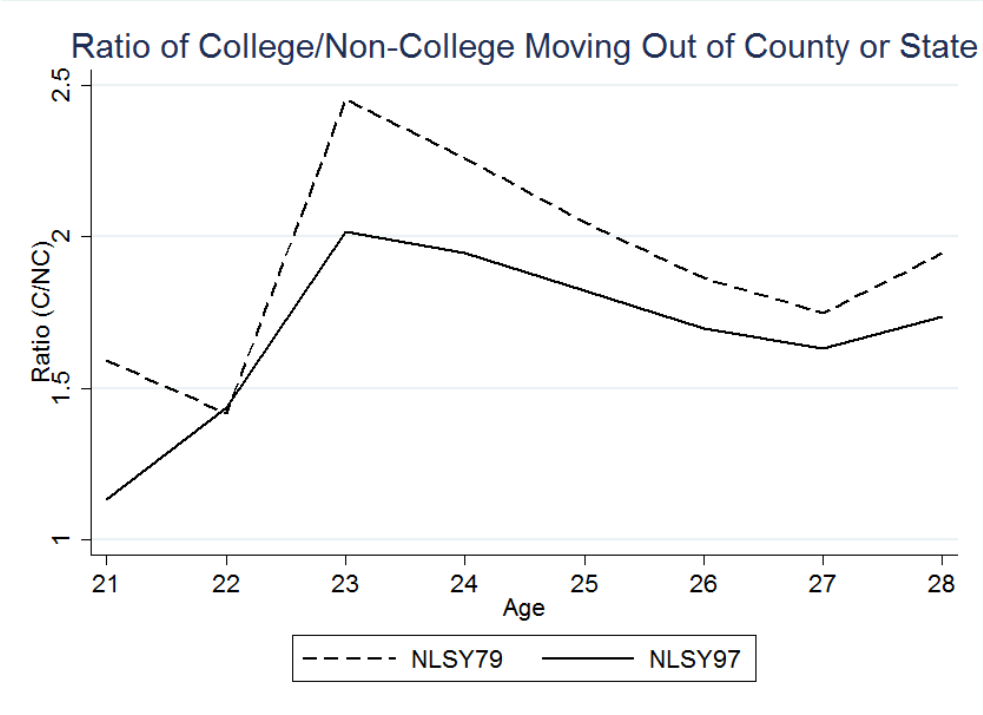
The ratio between the proportion of college and non-college graduates that moved out of county or state at each age is shown in Figure 2.4. The ratio at each age is defined as the proportion of college graduates that moved out of county or state divided by the proportion of non-college graduates that moved out of county or state. Hence, a ratio greater than one implies college graduates are more likely to move out of county or state than non-college graduates. For both the NLSY79 and NLSY97 cohorts, the ratio reaches a peak at the age of 23. This suggests that college and non-college graduates are most different in migration rates at that age. Further, the ratio is greater for the NLSY79 cohort than the NLSY97 cohort. This implies that college graduates in the NLSY79 cohort are more likely to move relative to their non-college graduate counterparts.

Table 2.2: Probability of Moving

	NLSY79		NLSY97	
	(1)	(2)	(1)	(2)
College Graduate	0.37***	0.46***	0.27***	0.45***
	(0.02)	(0.03)	(0.02)	(0.03)
Age	0.27***	-0.15**	0.52***	0.19***
	(0.05)	(0.06)	(0.06)	(0.08)
Age Squared	-0.006***	0.002*	-0.01***	-0.004***
	(0.0009)	(0.001)	(0.001)	(0.001)
One Child	-0.12***	-0.14***	-0.15***	-0.27***
	(0.02)	(0.03)	(0.03)	(0.04)
Two Children	-0.24***	-0.19***	-0.23***	-0.24***
	(0.03)	(0.04)	(0.04)	(0.05)
More than Two Children	-0.12***	-0.07	-0.23***	-0.24
	(0.05)	(0.06)	(0.04)	(0.07)
Real Hourly Wage	-0.0006	-0.0002	-0.0005	-0.002***
	(0.0005)	(0.0004)	(0.0005)	(0.0008)
Intercept	-4.37***	0.60	-7.63***	-3.84***
	(0.64)	(0.75)	(0.74)	(0.93)

*** denotes statistical significance at a 1% level and ** denotes significance at the 5% level. Standard errors are in parenthesis below regression coefficients. The dependent variable in regression (1) is whether there was a move out of county or state in that year, and in regression (2) it is whether a far move, defined as a move greater than 50 miles, occurred.

Figure 2.4: Ratio of College to Non-college Moving Out of County or State



2.3 Results

A fixed effects logit regression is used to analyze the impact of migration on cohabitation. Two different measures of migration are again used: (1) whether a move out of county or state happens in the next year and (2) whether a far move of over 50 miles occurs in the next year. The dependent variable is whether or not the individual is cohabiting. The left panel of Table 2.3 shows the result for the NLSY79 cohort and the right panel shows the result for the NLSY97 cohort.

The coefficients shown are in log odds. A negative coefficient implies a decrease in log odds of the dependent variable occurring. In both the NLSY79 and NLSY97 cohorts and for both measures of migration, moving next year decreases the odds of cohabiting in the current year. However, the result is only statistically significant for the NLSY97 cohort.

Table 2.3: Odds of Cohabiting

	NLSY79		NLSY97	
	(1)	(2)	(1)	(2)
Moving Out of County/State Next Year	-0.10		-0.42***	
	(0.07)		(0.04)	
Far Move Occurring Next Year		-0.09		-1.08***
		(0.09)		(0.07)
Age	1.35***	1.34***	1.23***	1.23***
	(0.13)	(0.13)	(0.11)	(0.11)
Age Squared	-0.02***	-0.02***	-0.02***	-0.02***
	(0.002)	(0.002)	(0.002)	(0.002)
One Child	-0.10	-0.10	0.18***	0.17***
	(0.07)	(0.07)	(0.06)	(0.06)
Two Children	-0.58***	-0.58***	-0.61***	-0.60***
	(0.10)	(0.10)	(0.08)	(0.08)
More than Two Children	-0.43***	-0.43***	-0.96***	-0.98***
	(0.14)	(0.14)	(0.13)	(0.13)
Real Hourly Wage	-0.001	-0.001	-0.0003	-0.0003
	(0.001)	(0.001)	(0.001)	(0.001)

*** denotes statistical significance at a 1% level and ** denotes significance at the 5% level. Standard errors are in parenthesis below regression coefficients. Controlled for individual fixed effects.

2.4 Discussion

The above analysis shows that having a big move in the future could impact the probability of whether an individual chooses to cohabit. College graduates are more likely to move and less likely to cohabit compared to non-college graduates, which is consistent with the hypothesis of this paper. Moreover, the difference between college and non-college migration and cohabitation rates are also consistent with my hypothesis. Recall from Figure 2.4 that 23 is the age at which college and non-college graduates

have the most dissimilar migration rates. This should imply that the age at which cohabitation rates are most dissimilar is right before age 23, which Figure 2.2 shows. This is reinforced by Table 2.3, where a big future move lowers the odds of cohabitation occurring.

While this explains the cross-sectional difference between college and non-college graduates, it does not address the difference between the 1979 and 1997 cohorts. Recall from Figure 2.1 that there is an increase in cohabitation rate for both college and non-college graduates from the NLSY79 to NLSY97 cohort. In order for migration to be a possible factor in the increase in cohabitation rates between the 1979 and 1997 cohorts, the 1979 cohort must have a higher probability of moving. However, this is not true from the NLSY data (as seen in Figure 2.3). This is different from what other papers in the migration literature have found. For example, Molloy et al. (2011) and Kaplan and Schulhofer-Wohl (2015) document that interstate migration has declined between 1980 and 2010, and that migration out of county followed a similar pattern.

To make a comparison, I use the Census Bureau's Current Population Survey, used by both Molloy et al. (2011) and Kaplan and Schulhofer-Wohl (2015), but focus on ages 21 to 28. Table 2.4 shows the migration rate out of county, out of state, and out of county or state for those ages 21 to 28 for non-college and college graduates. Migration rate is shown in two different time periods, 1981-1992 and 2001-2012, which correspond to the periods when the NLSY79 and NLSY97 cohorts, respectively, would have been between ages 21 and 28[2]. The CPS data confirms that migration rates should be lower for the NLSY97 cohort compared to the NLSY79 cohort.

[2] For the NLSY79 cohort, the first year where an individual can be 21 is 1978. However, lack of data before 1981 is the reason why 1981 was chosen as the start year.

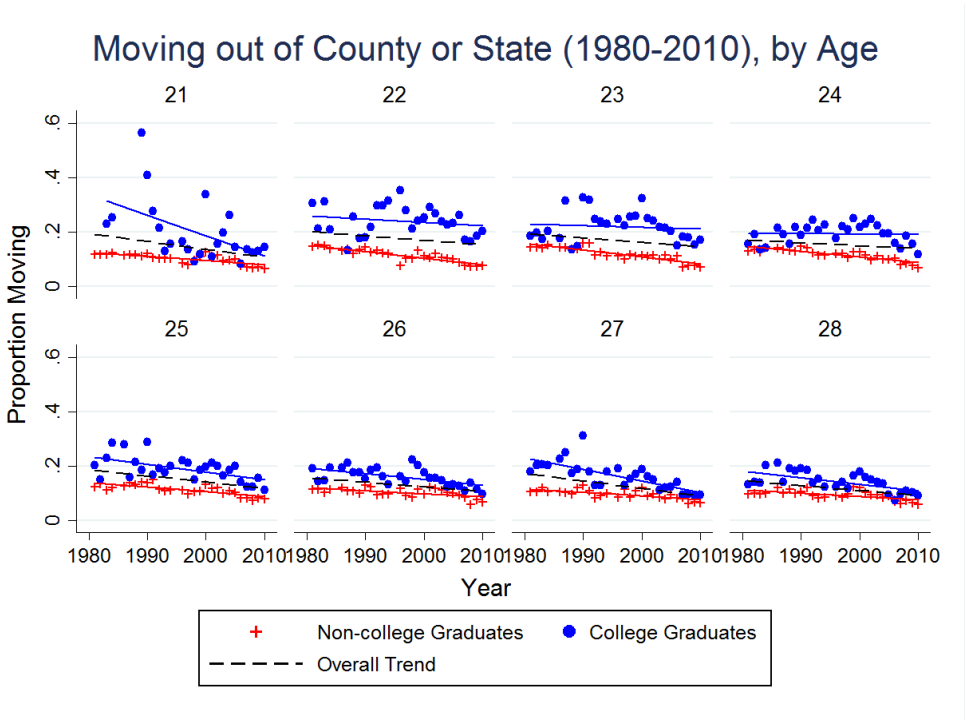
Table 2.4: Migration Rate Out of County, State, or Either (Percentages)

	1981-1992	2001-2012
Non-College Graduates		
Out of County	6.4	5.0
Out of State	4.8	3.5
Out of County or State	11.2	8.5
College Graduates		
Out of County	9.1	7.4
Out of State	8.8	7.1
Out of County or State	17.9	14.6

I also present the migration rate for each age between 21 and 28 by education and by year in Figure 2.5. An important caveat of doing this, however, is that the CPS does not have very many observations in the early 1980s when categorized so finely. This can be seen in some panels of Figure 2.5 with unrealistically high migration rates. Again, the overall trend tends to be declining from 1980 to 2010 for both non-college and college graduates.

The CPS data thus seems to contradicts the NLSY data in how migration rate changed between the NLSY79 and NLSY97 cohorts. According to the CPS data, the decrease in migration rate could have contributed to the increase over time in cohabitation rates for both college and non-college graduates.

Figure 2.5: Moving Out of County or State (CPS data)



2.5 Conclusion

In this paper, I analyze whether and how out of county or state migrations impact cohabitation. I study migration and cohabitation behaviors of non-college and college graduates in the NLSY79 and NLSY97 cohorts. Cross-sectionally, I document that non-college graduates and college graduates in both cohorts are most different in whether they cohabit about one year before they are most different in whether they move. I also show that college graduates are more likely to migrate, and migration in the future decreases the odds of cohabiting. Thus, I suggest that a higher possibility of a future migration could have a deterring effect on cohabitation because cohabitation is a form of settling down. These results suggest that migration could explain the cross sectional differences between non-college and college graduates.

Over time, cohabitation rate has increased and migration rate has decreased for both non-college and college graduates. Using the Census Bureau's Current Population

Survey, I show that migration rate has fallen between the two cohorts for both education groups. Lower migration rate implies a higher cohabitation rate, which is consistent with data. I also document that the two education groups are more similar in both migration and cohabitation in the NLSY97 than in the NLSY79. This suggests that migration can also be a factor in explaining the difference over time in non-college and college cohabitation rate.

This paper suggests that future research in how different groups cohabit should take into account migration as a factor. While beyond the scope of this paper, a careful estimate of the magnitude of migration's impact on cohabitation is an important topic for future research.

References

Adamopoulou, E. (2010): "Will You "Quasi-marry" Me? The Rise of Cohabitation and Decline of Marriages," *Universidad Carlows III de Madrid Working Paper*.

Aitchison, J. (1986): *The Statistical Analysis of Compositional Data*, London: Chapman & Hall.

Becker, G. (1973): "A Theory of Marriage: Part I," *The Journal of Political Economy*, Volume 81, 813–846.

——— (1974): "A Theory of Marriage," *NBER Economics of the Family: Marriage, Children, and Human Capital*, 299–351.

——— (1981): *A Treatise on the Family*, Cambridge: Harvard University Press.

Blau, D. and W. van der Klaauw (2010): "What Determines Family Structure?" *Institute for the Study of Labor (IZA) Discussion Paper Series*.

Brien, M., L. Lillard, and S. Stern (2006): "Cohabitation, Marriage, and Divorce in a Model of Match Quality," *International Economic Review*.

Bumpass, L. and H.-H. Lu (2000): "Trends in Cohabitation and Implications for Children's Family Contexts in the United States," *Population Studies*, 29–41.

Bumpass, L. and J. Sweet (1995): "Cohabitation, Marriage and Union Stability: Preliminary Findings from NSFH2," *NSFH Working Paper No. 65*.

Chade, H. and G. Ventura (2005): "Income Taxation and Marital Decisions," *Review of Economic Dynamics*, 565–599.

Choo, E. and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, Vol. 114, 175–201.

Fields, J. and L. Casper (2001): "America's Families and Living Arrangements," Tech. rep.

Finer, L. and S. Henshaw (2006): "Disparities in Rates of Unintended Pregnancy in the United States, 1994 and 2001," *Perspectives on Sexual and Reproductive Health*, 90–96.

Gaertner, L. and V. Foshee (1999): "Commitment and the perpetration of relationship violence," *Personal Relationships*, 227–239.

Gemici, A. and S. Laufer (2014): "Marriage and Cohabitation," *Working Paper*.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2012): "Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment and Married Female Labor-Force Participation," *University of Pennsylvania Population Studies Center Working Paper Series*.

Johnson, M. and K. Ferraro (2000): "Research on Domestic Violence in the 1990s: Making Distinctions," *Journal of Marriage and the Family*, 948–963.

Kambourov, G. and I. Manovskii (2009): "Accounting for the Changing Life-Cycle Profile of Earnings," *Working Paper*.

Kaplan, G. and S. Schulhofer-Wohl (2015): "Understanding the Long-Run Decline in Interstate Migration," *Federal Reserve Bank of Minneapolis Working Paper 697*.

Lichter, D. (2012): *Childbearing Among Cohabiting Women: Race, Pregnancy, and Union Transitions*, 209–219.

Lichter, D., Z. Qian, and L. Mellott (2006): "Marriage or Dissolution? Union Transitions Among Poor Cohabiting Women," *Demography*, 223–240.

Light, A. and Y. Omori (2008): "Economic Incentives and Family Formation," *Working Paper*.

Lillard, L., M. Brien, and L. Waite (1995): "Premarital Cohabitation and Subsequent Marital Dissolution: A Matter of Self-Selection," *Demography*.

Lundberg, S. and R. Pollak (2013): "Cohabitation and the Uneven Retreat from Marriage in the U.S., 1950-2010," *NBER Working Paper 19413*.

Martinez, G. (2015): "Three Decades of Nonmarital First Births Among Fathers Aged 15-44 in the United States," Tech. rep., Hyattsville, MD.

Molloy, R., C. Smith, and A. Wozniak (2011): "Internal Migration in the United States," *Journal of Economic Perspective*, 172–196.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015): "Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]," Minneapolis: University of Minnesota.

Smith, L. (2006): "The Marriage Model with Search Frictions," *Journal of Political Economy*, Volume 114.

Smock, P. and W. Manning (2004): "Living Together Unmarried in the United States: Demographic Perspectives and Implications for Family Policy," Tech. rep.

Stevenson, B. and J. Wolfers (2007): "Marriage and Divorce: Changes and their Driving Forces," *Journal of Economic Perspectives*, 27–52.

A.1 Data Appendix

A.1.1 Relationship Time

A relationship time is constructed for each relationship, where the start of a relationship is defined as time 0. I then construct time -1, -2, -3, and so on to denote time before the start of the relationship and time 1, 2, 3, and so on to denote the time after the start of the relationship.

Table A.1 describes how relationship time is defined for each type of relationship. For those who cohabit and eventually marry, I record both the start of cohabitation and the start of marriage and the time before and after. This will allow me to better analyze the characteristics and behavior of those who cohabit and then eventually marry.

Not all relationships meet the conditions necessary to have relationship time go from -12 to 18. For example, a respondent may only be observed to be single starting 5 months before his relationship begins, and after cohabiting at time 0, he may get married 8 months later, and I only observe him to be married for 15 months before the last observation from the panel. This is how relationship time is constructed for this respondent: For this respondent's cohabiting relationship, the time will only run from -5 to 7, since I only observe him to be single five months before the start of his cohabitation, and the last month that he cohabits is seven months after the co-residential relationship starts. Then, time 0 is assigned when he gets married eight months after the start of cohabitation, and time 1, 2, 3,...,15 indicates one, two, and three,. . . ,fifteen months after the start of his marriage. Importantly, even though the respondent is with the same partner in cohabitation and marriage, I account for both the start of cohabitation and the start of marriage as time 0, which will allow me to see if labor characteristics differ at the start of cohabitation or marriage. Relationships where I observe cohabitation and marriage are not double counted.

Table A.1: Construction of Relationship Time

	Cohabit	Cohabit and Marry	Cohabit and Marry	Marry
Pre-relationship (t= -12, -11, -10,. . . , -1)	Must be either single and never married, or legally separated, divorced, or widowed and not in a cohabiting relationship	Must be either single and never married, or legally separated, divorced, or widowed and not cohabiting	Must be cohabiting with partner that respondent will eventually marry	Must be either single and never married, or legally separated, divorced, or widowed and not cohabiting
Start of relationship (t=0)	The first month where respondent is observed to be cohabiting with partner	The first month where respondent is observed to be cohabiting with partner	The first month where respondent is observed to be married to cohabiting partner	The first month where respondent is observed to be married to spouse
Relationship (t=1,2,3,. . . ,18)	Months after the start of the relationship that the respondent is observed to be cohabiting with partner. Relationship must not transition into marriage.	Months after the start of the relationship that the respondent is observed to be cohabiting (and not married to) with partner.	Months after the start of the relationship that the respondent is observed to be married to spouse.	Months after the start of the relationship that the respondent is observed to be married to spouse.

A.1.2 Education

For each respondent not in school and for each month m, I categorized their education into one of two groups: college graduates and non-college graduates. College graduates are those who have at least a bachelor's degree and are not currently in school (including graduate school) in month m, and non-college graduates are those who meet one of either two conditions: 1) not observed to obtain a bachelor's degree in the sample or 2) those who are observed to obtain a bachelor's degree at $m^\star > m$ and are not currently in school.

I choose not to include months where a respondent is in school because being in school changes the incentives of labor market work. Not counting those who are in school, about 81% of the sample is non-college graduates and 19% are college graduates in 2011 (the last year with complete data in the panel).

Average age at the start of each type of relationship by education level can be seen in Table A.2. For those who cohabit and eventually marry their partners, the age at the start of cohabitation and the age at the start of marriage are both shown.

Table A.2: Average Age at Start of Relationship, by Education

	Non-College	College
Cohabit only	21.6	26.1
Cohabit and Marry	21.8	25.3
Cohabit and Marry	23.8	26.6
Marry only	22.7	25.7

A disadvantage of the NLSY97 (up to round 15) in analyzing co-residential relationship by education groups is that relationships of college graduates may be more noisy than for non-college graduates, since the average age at the beginning of a relationship is closer to the maximum age in the sample. With each additional round of the NLSY97, there will be a fuller set of data for relationships of college graduate respondents. While the NLSY97 does not provide education information for co-residential partners directly, respondents are asked to list members of their household and their relationship to him or her. By exploiting the household roster and linking each partner to their spot on the household roster, data on partner education, employment status, and income was

collected.

However, not every respondent reported the education level of everyone in their household, thus some co-residential partners had no education level observed. The number of each relationship groups where partner's education is observed is summarized in Table A.3.

Table A.3: Relationship Groups with Observed Partner's Education

	Cohabit	Cohabit and Marry	Marry	Total
Partner is Non-College Grad	1176 (43.2%)	1063 (39.1%)	481 (17.7%)	2720 (100%)
Partner is College Grad	147 (28.6%)	225 (43.8%)	142 (27.6%)	514 (100%)
Total	1323	1288	623	4193

From respondent and partner education, we can calculate the degree of assortative mating at the beginning of a relationship (time 0). For those who cohabit and then marry, the beginning of the relationship is the start of the cohabiting phase. A summary is shown in Table A.4.

Table A.4: Assortative Mating by Education and Relationship Group

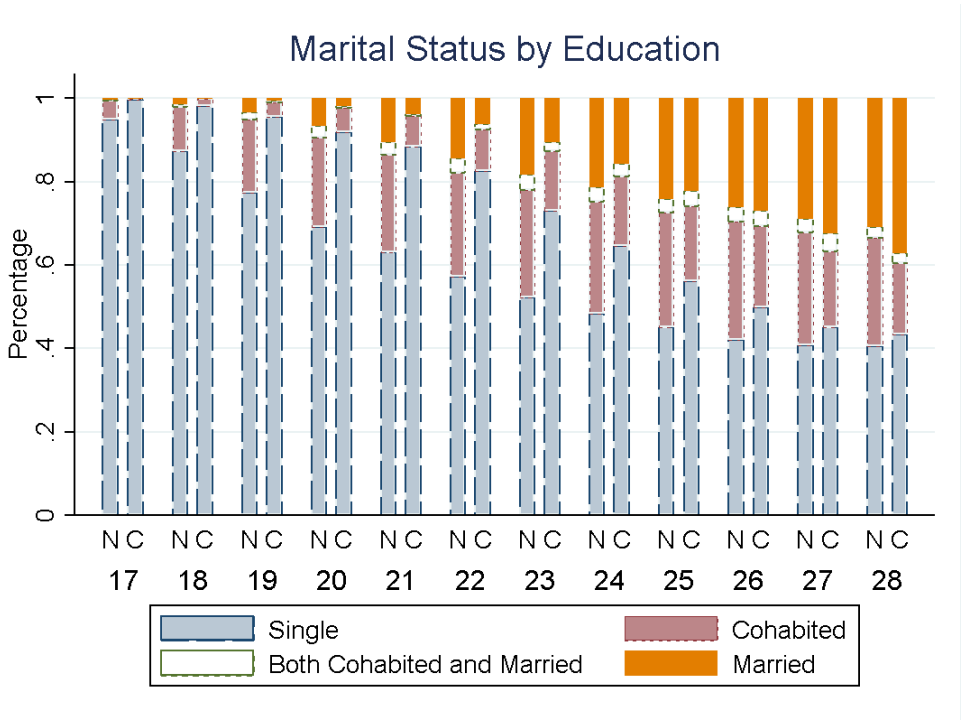
	Non-College	College
Cohabit, same education	44.7%	14.5%
Cohabit, different education	2.4%	14.1%
Cohabit and Marry, same education	31.9%	26.3%
Cohabit and Marry, different education	2.4%	17.1%
Marry, same education	16.9%	18.5%
Marry, different education	1.6%	9.5%

The first observation is that college graduates are more likely to be in a co-residential relationship with someone who does not have a college education. This can be seen in the percentage of college graduates who cohabit with those with different education (14.1% compared to 2.4% for non-college grads), cohabit and marry (17.1% compared to 2.4% for non-college grads), and marry (9.5% compared to 1.6%).

Next, while 47.1% of non-college grads are in cohabiting relationship that did not end in marriage or where marriage has not happened, 28.6% of college grads are in that category. The flip side of this observation is that 71.4% of college graduates' relationships would transition into marriage while 52.9% of non-college graduates' relationships would.

Lastly, college grads are more likely to stay single at each age, as seen in Figure A.1. The cohabited and married group in the figure below means the percentage of respondents who were observed to have both cohabited and married at a specific age. This is not the same as those who married from cohabitation. The cohabit and eventually married group is included as part of cohabitation if the respondent at a certain age is in the cohabitation phase, and included as part of married if the respondent at a certain age is in the marriage phase.

Figure A.1: College graduates more likely to remain single



A.1.3 Labor Supply

Employment data is provided for each job that a respondent held prior to the annual interview date. Unique employer id can be used to match jobs from previous years, and an employment history can be constructed for each respondent. Further, hours worked per week, work schedule, hourly wage, and other employer-specific characteristics can be matched with each job. For months where the respondent is not observed to be with an employer, I use employment status data for each individual to determine whether the respondent is in the labor force or unemployed.

Employment status in the NLSY97 is found at the weekly level, coded as shown in Table A.5.

Table A.5: Employment Status Coding

Employment Status	Description
0	No information reported to account for week, job dates indeterminate
1	Not associated with an employer, not actively searching for an employer job
2	Not working (unemployment vs. out of labor force cannot be determined)
3	Associated with an employer, periods not working for the employer are missing
4	Unemployed
5	Out of labor force
6	Active military force
9701-201399	Unique employer ID (for those who are associated with an employer and periods working for employer are observed)

To convert weekly employment status into a summary monthly employment status, I first assigned specific weeks into a specific month, as shown in Table A.6.

Table A.6: Assignment of Weeks to Corresponding Month

Month	Weeks	Month	Weeks
1	1-4	7	27-30
2	5-8	8	31-34
3	9-13	9	35-39
4	14-17	10	40-43
5	18-21	11	44-48
6	22-26	12	49-52

Then, I took the maximum over employment status for the weeks assigned to each

month. Taking the maximum over the employment status variable meant that a respondent could have been associated with an employer for one week out of four, and I will count that agent as having been employed in that month.

A.1.4 Labor Force Participation

Labor force status at the beginning of a relationship is shown in Table A.7 for college and non-college men and women. College men and women are more likely to be in the labor force compared with non-college men and non-college women. Men and women without a college degree are more likely to be in the labor force if they cohabit and then marry.

Table A.7: Labor Force Participation

	Cohabit	Cohabit and Marry	Marry
Non-college men	85.7%	92.5%	90.4%
College men	97.7%	95.0%	94.0%
Non-college women	77.6%	84.0%	77.0%
College women	95.0%	94.5%	92.2%

A.1.5 Employment

Employment at the beginning of a relationship is shown in Table A.8 for college and non-college men and women. If the respondent is associated with an employer, then they are counted as employed. Only those who are listed specifically as unemployed in the respondent's employment status are counted as unemployed. I do not count those who are not working and where unemployment and not being in the labor force cannot be determined.

Table A.8: Employment

	Cohabit	Cohabit and Marry	Marry
Non-college men	89.5%	94.7%	93.7%
College men	98.6%	97.9%	98.7%
Non-college women	90.9%	94.9%	90.8%
College women	95.3%	97.3%	93.2%

A.1.6 Hours Worked

Hours worked is the sum of hours of first six jobs listed. Number of hours worked can be zero if person is associated with an employer but had zero hours worked that week. Number of hours worked is missing if there are no associated employers and the number of hours reported is zero or missing. The total number of employed workers (used to find the average hours worked per employed worker by relationship groups) is then the total hours worked divided by the total number of workers who had a non-missing hours worked.

A.1.7 Income

Income is the sum of the respondent's hourly wage times hours worked at that hourly wage from up to his first six concurrent jobs recorded. An individual has more than one concurrent job if they report more than one employer and have positive hours worked for more than one job. An insignificant portion of the NLSY97 data had weeks where the respondent held more than six jobs. The wage for each job was first put into 2005 dollars, and then multiplied by hours worked at that job to get real weekly income from that specific job. Summing up across up to six jobs gives the individual's real weekly income. Weekly income is defined as:

$$I = \sum_{j=1}^{6} w_j h_j$$

For those with less than six concurrent jobs, total weekly income is the sum of income from all jobs. There are some observations that there are assumed to have measurement

error. For example, some reported an hourly wage of $45,000 for a specific employer or more than 168 hours a week working at a specific job. Outliers are dropped before computing income. For hours worked, an outlier is greater than 100 hours worked per week (about 14 hours a day for 7 days). For hourly pay, an outlier is considered $500 per hour. Both hours and pay outliers are well above the 99th percentile of data.

A.1.8 Dating

Yearly number of dates gone on by each youth is recorded between the years 2002-2008, when interviewed youths were between ages 23 and 29. To translate dating into effort before the start of a relationship, I look at the number of dates that the respondent said he or she had gone on in the previous year, starting from 6 months before his or her co-residential relationship starts to 6 months after. While not an exact way of measuring how many dates the respondent went on with their partner before the start of their co-residential relationship, it serves as a rough estimate of effort. For example, someone who said they went on 5 dates in the year before cohabiting is said to have put in less effort than someone who answered they had gone on 50 dates in the previous year.

While it is possible to split the data into both education and relationship status, because marriage and cohabitation into marriage happens later in life for college graduates, the number of observations for only years 2002-2008 becomes somewhat of an issue. Still, it can be seen that those who marry will have more dates before the start of a relationship, and those who cohabit and those who cohabit and eventually marry will also act similar, as seen in Table A.9.

Table A.9: Number of Dates before Relationship, by Education

	Non-College	College
Cohabit	25.23	43.28
Cohabit and Marry	23.08	42.62
Marry	36.41	47.27

A.1.9 Robustness Check

One concern of the data may be the right censoring of those in the cohabit only group, as that group may contain those who will marry eventually but marriage is not observed yet given data limitations. I predict that this should have the greatest impact on college graduates, since their relationships start later on average than non-college graduates, and thus more college graduate relationships are prone to being right censored. I outline how many relationships fall into this case in Table A.10, and construct a new cohabit only group based on the criteria that an end to the relationship was observed. I show that the results in this paper do not change for the most part.

Table A.10: Number of Relationships by Relationship Groups and Education Level

	Non-College	College	Total
Cohabit (no removal of censored relationships)	3022	375	3397
Cohabit (with removal of censored relationships)	1893	149	2042
Percentage change in number of observations	-37%	-60%	-40%

Hours worked for the cohabit and marry group remains higher at the start of the relationship, compared to the group that only cohabits, as seen in Table A.11.

Figure A.2 shows that the results from Figure 1.1 are the same for non-college graduate men and women. College-graduated men who cohabit and later marry still work more than those who only cohabit, but those who cohabit only now work more than those who marry directly. College graduated women who cohabit now work more than those who will eventually marry from cohabitation.

Two important points should be considered with these new results. 1) The 60% loss in observations for college cohabitors will have a selection issue. Because of the later average age of cohabiting for college graduates, and because the last year where there is a full set of observations for the entire sample is 27, the college graduated cohabitors who already were cohabiting and separated are likely different than the average college graduate, given the narrow time window. 2) These results are now based on only 149 relationships, and thus it can be expected that there will be an increased amount of noise in the results. Importantly, for non-college graduates where I do have sufficient

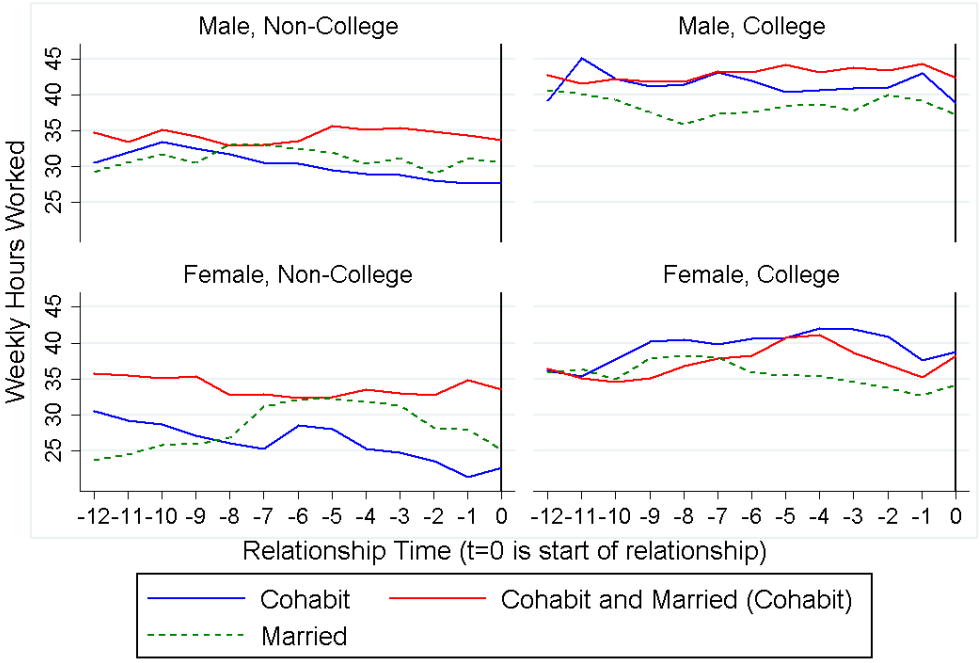
observations, the result is identical.

Table A.11: Weekly Hours Worked at Start of Relationship, Robustness Check

	Coefficient	(Std. Error)
Reference: Non-college & Cohabit		
Non-college & Cohabit and Marry	1.75***	(0.60)
Non-college & Marry only	1.40**	(0.70)
College & Cohabit only	4.79***	(1.19)
College & Cohabit and Marry	5.95***	(0.94)
College & Marry only	4.68***	(1.13)
Female	-4.31***	(0.48)
Age	5.19**	(2.02)
Age Squared	-0.10**	(0.04)
One Child	0.23	(1.65)
Two Children	1.43	(2.51)
More than Two Children	2.39	(4.44)
Real Hourly Wage	-0.08***	(.01)
Intercept	-27.51	(24.92)

*** denotes statistical significance at a 1% level and ** denotes significance at the 5% level.

Figure A.2: Weekly Hours Worked Before the Start of Relationship, Robustness Check



A.2 Model Appendix

A.2.1 Continuity of the value of cohabitation and marriage

The value of cohabitation and marriage are

$$\text{v(cohab)} = \begin{cases} -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \ln(w_i(j_x)h) & \theta < 0 \\ -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta \\ \quad + \beta \left[\theta \left(\ln(w_i(j_x)h + I^\star) - p j_x h^\alpha + \frac{1}{2}\theta \right) \right. & 0 \leqslant \theta \leqslant 1 \\ \quad \left. + (1-\theta) \ln(w_i(j_x)h) \right] & \\ -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta \\ \quad + \beta \left[\ln(w_i(j_x)h + I^\star) + \frac{1}{2}\theta - p j_x h^\alpha \right] & \theta > 1 \end{cases}$$

$$\text{v(marry)} = \begin{cases} -p j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta \ln(w_i(j_x)h) & \theta < 0 \\ -p j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta \\ \quad + \beta \left[\theta \left(\ln(w_i(j_x)h + I^\star) + \frac{1}{2}\theta \right) \right. & 0 \leqslant \theta \leqslant 1 \\ \quad \left. + (1-\theta) \ln(w_i(j_x)h) \right] & \\ -p j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta \\ \quad + \beta \left[\ln(w_i(j_x)h + I^\star) + \frac{1}{2}\theta \right] & \theta > 1 \end{cases}$$

It can be seen that both the value of cohabitation and marriage are continuous for $\theta < 0$, $\theta > 1$, and $0 < \theta < 1$. It needs to be shown that $\lim_{\theta \to 0^+} v(\text{cohab}) = v(\text{cohab}|\theta = 0) = \lim_{\theta \to 0^-} v(\text{cohab})$.

$$\begin{aligned}
&\lim_{\theta\to 0^+} -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta + \beta\ln(w_i(j_x)h) \\
&\quad = -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \beta\ln(w_i(j_x)h) \\
&\quad = \lim_{\theta\to 0^-} -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \theta \\
&\quad + \beta\left[\theta\left(\ln(w_i(j_x)h + I^\star) - pj_x h^\alpha + \frac{1}{2}\theta\right) + (1-\theta)\ln(w_i(j_x)h)\right]
\end{aligned}$$

Also, $v(\text{cohab}|\theta = 0) = -j_x h^\alpha + \ln(w_i(j_x)h + I^\star) + \beta\ln(w_i(j_x)h)$. Thus, the value of cohabiting is continuous. The same is true for the value of marriage.

A.2.2 Value of marriage increasing at a greater rate in θ than the value of cohabitation for $0 \leq \theta \leq 1$

When $0 \leqslant \theta \leqslant 1$, the value of cohabitation and the value of marriage can be written in the following way

$$\begin{aligned}
v(\text{cohab}|0 \leqslant \theta \leqslant 1) &= (1 - \beta p j_x h^\alpha - \beta\ln(w_i(j_x)h) + \beta\ln(w_i(j_x)h + I^\star))\,\theta + \frac{\beta\theta^2}{2} \\
&+ [-j_x h^\alpha + \beta\ln(w_i(j_x)h) + \ln(w_i(j_x)h + I^\star)]
\end{aligned}$$

$$\begin{aligned}
v(\text{mar}|0 \leqslant \theta \leqslant 1) &= (1 - \beta\ln(w_i(j_x)h) + \beta\ln(w_i(j_x)h + I^\star))\,\theta + \frac{\beta\theta^2}{2} \\
&+ [-pj_x h^\alpha + \beta\ln(w_i(j_x)h) + \ln(w_i(j_x)h + I^\star)]
\end{aligned}$$

It can then be seen that because $\beta p j_x h^\alpha > 0$, the slope of the value of marriage increases at a greater rate with θ.