Credit Scores and Committed Relationships

November 26, 2023

Abstract

We document a sizable positive correlation in the two individuals' credit scores at the time when they form a committed relationship, controlling for similarities in their socioeconomic and demographic characteristics. Moreover, couples with lower average credit scores and larger score differentials at relationship-onset are more likely to separate in the future. Match quality in credit scores predicts separations over and above its association with subsequent financial distress and limited credit access, which themselves can influence relationship outcomes. We explore possible behavioral traits reflected in credit scores that may help account for their predictability regarding relationship dissolutions.

Keywords: Credit scores, Committed relationships, Household finance, Trustworthiness

JEL codes: D14, G21, J12

1 Introduction

In today's economy, credit scores have become omnipresent in household life. Used in underwriting most household debt, credit scores are a foremost factor that plays a critical role on consumers' access to credit and borrowing costs, which in turn affect their capability of purchasing durable goods and smoothing consumption through income fluctuations. Outside the credit market, credit scores are increasingly used in other areas, such as the auto insurance and rental housing markets. In particular, survey evidence suggests that up to 60 percent of employers run credit checks on applicants as part of the hiring process.

Motivated by the growing prominence of credit scores in everyday life, we explore their relevance in partner selection and relationship dissolution using a large, proprietary data set. Our results point to a quantitatively large and statistically significant association between credit scores and both the formation and dissolution of committed relationships (marriages and long-term cohabitations). Specifically, we document three empirical patterns. First, credit scores are positively correlated with the likelihood of forming a committed relationship. Second, people tend to form relationships with partners of like credit scores, even after taking into account other socioeconomic and demographic similarities. In particular, we note that the credit score similarity does not merely reflect similarities in income and earning capacity. Third, in spite of this positive correlation, within-couple differences in credit scores at the onset of relationships exist and are highly predictive of subsequent separations. Such a predictive power holds over and above credit scores' association with future conditions of financial distress and credit access, which themselves can influence relationship outcomes.

While our research design does not fully establish a causal effect, the fact that initial credit score conditions predict relationship outcomes is a new contribution to the literature of relationship dynamics. In addition, our statistical results are estimated from committed relationships formed within a large, near 5 percent random sample of U.S. adult population and are thus likely generalizable to the broad population. Broadly speaking, our study contributes to a range of topics related to (i) how individuals sort into committed relationships, (ii) how match quality help predict relationship outcomes, and (iii) the broad interpretation of credit scores and its relevance beyond predicting debt defaults.

On the first topic, our results suggest credit scores as a new characteristic by which individuals sort themselves into committed relationships, complementing a voluminous literature that focuses on race, educational attainment, earnings capacity, parental wealth, social caste, and physical appearance.[1]

On the second topic, in spite of this voluminous literature on assortative mating, the empirical evidence on how match quality between spouses may lead to relationship dissolution appears to be more scant (Lyngstad and Jalovaara, 2010; Mortelmans, ed, 2020). In the framework of Becker et al. (1977), relationship dissolution is largely unpredictable and often owes to shocks that transpire *during* the relationship (Weiss and Willis, 1997). Work searching for ex ante factors that predict separations often takes place in marriage clinics (Gottman, 2014). This paper helps bridge this gap in the literature and documents that relationship outcome can be predicted using *initial* or even pre-relationship credit score match quality. In particular, our analysis underscores that the mismatch in credit scores may limit household credit use during the relationship and heighten financial vulnerability, thereby undermining relationship durability.

On the third topic, our paper belongs to a nascent literature that seeks for the broader information contents of credit scores and their implications outside credit markets. Notable examples include Chen et al. (2013) and Bos et al. (2018), who focus on the labor market. Because the credit score conditions at the onset of the relationships are strong predictors of future separations beyond the credit market channels, we consider the possible behavioral traits reflected in credit scores that may also have a bearing on relationship outcomes. Survey-based analysis and historical anecdotes corroborate our conjecture that credit scores reflect individu-

[1] See, for example, Garfinkel et al. (2002); Blackwell and Lichter (2004); Watson et al. (2004); Chiappori et al. (2012); Banerjee et al. (2013); Charles et al. (2013).

als' underlying trustworthiness that influences the willingness to honor financial and personal commitments.

Sociologists and demographers have explored the link between personal traits and relationship dissolutions, often along the so-called "big-five" traits (Solomon and Jackson, 2014; Arpino et al., 2022). Consistent with our conjecture, Boertien et al. (2015) show that high levels of "conscientiousness," one of the big-five, strongly predict relationship durability.[2] Indeed, because household relationships are subject to fewer formal, contractual restraints and largely rely on informal, difficult-to-enforce norms, personal traits linked to trustworthiness and conscientiousness are all the more important in maintaining such relationships (Arrow, 1972; Weiss, 1997; Guiso et al., 2004; Karlan, 2005). Thus, one contribution of our paper to this research agenda is to introduce credit scores, available in high-quality administrative data, as a possible indicator for otherwise hard-to-measure personal traits.

Our data track individual consumers over time, offering a rare opportunity of observing credit scores before and after relationship formation. Thus, a unique strength of our research design is to relate credit score conditions at the time of (and prior to) relationship formation with subsequent relationship status changes, thereby circumventing two important endogeneity concerns. First, spouses' credit scores tend to comove and converge during the course of the relationship. Second, couples's financial decisions tend to influence both their credit scores and relationship prospects.

A limitation of our analysis is the lack of detailed individual-level demographic and income data as control variables. To the extent these factors may be correlated with credit scores and also have an effect on relationship formation and dissolution, our estimates may be biased because of omitted variables. We mitigate this limitation by using census block group-level statistics as proxies for individual characteristics. We leverage the Panel Study of Income Dynamics data to show that, for sufficiently small geographies, such proxies perform well in reducing the omitted variable biases.

[2]Other personal traits of the big-five include agreeableness, extraversion, neuroticism, and openness.

Furthermore, we argue that not observing individual income is not likely to qualitatively bias our main results for the following reasons. First, essentially all credit scoring models exclude income from their algorithms, suggesting a limited statistical merit of income in predicting credit scores, conditional on the variables included in these algorithms. Using a random sample where individual characteristics and credit scores are both contemporaneously observed, Beer et al. (2018) find only a moderate correlation (0.3–0.4) between income and credit scores. Similarly, other characteristics such as race and education has little incremental power accounting for credit score variations. Second, we find that the within-couple correlation in credit scores is substantially larger than that observed in partners' income (0.63 versus 0.38) estimated in household survey data, reassuring a limited income-credit score linkage. That said, research on the relationship between credit scores and income has been inconclusive, with some analysis indicating a stronger correlation (see, for example, Albanesi et al. (2020)) and calling for more research using better data.

Finally, we note that we remain agnostic regarding the extent to which partners are aware of each other's credit scores prior to and during the relationship. With the increasing awareness of their importance in personal finance, we expect credit scores to become more visible between partners over time, and there has been extensive media attention on the importance of credit scores in relationship building and maintenance (Silver-Greenberg, 2012; Ettin, 2013).

The remainder of the paper is organized as follows. Section 2 describes the data used in this paper and section 3 introduces the algorithm we use to identify relationship formation. Sections 4 introduces the econometric framework, with a discussion on the statistical soundness of the proposed proxies and the omitted variable-bias likely being limited. Section 5 presents the main results, and Section 6 explores various channels in which credit scores may influence relationship outcomes. Section 7 concludes and sets an agenda for future research.

2 Data Description

2.1 The Consumer Credit Panel/Equifax Data

For individual-level data on credit scores and the formation and dissolution of committed relationships, we turn to the Federal Reserve Bank of New York's Consumer Credit Panel/Equifax (henceforth the CCP). Equifax is one of the three largest U.S. credit reporting agencies. These data are extensively used to analyze how households' liabilities evolve over time. Beginning in the first quarter of 1999, the data are a quarterly panel that tracks a 5 percent random sample (the "primary sample") of all U.S. adults with a valid credit history.[3] Our sample ends in the second quarter of 2014 and covers a total of 62 quarters. A unique feature of the CCP is its large sample size, with the primary sample containing about 12 million consumers in a typical quarter, which is crucial for our research design. All personally identifiable information is removed by Equifax.

In addition to the primary sample, the CCP also follows consumers who live at the same address as a consumer from the primary sample. However, these non-primary sample consumers are followed only for the duration they share the same address. Once they stop living with primary sample consumers, non-primary consumers and their credit records are no longer observed. Their information is also not observed before they live with primary sample consumers. Roughly 30 million non-primary sample consumers live with someone in the primary sample in a typical quarter.

In addition to a credit score (known as the Equifax Risk Score), variables in the CCP include: a information on loan balances; the delinquency status of various types of debt; bankruptcy, foreclosure, and other derogatory flags; the number of inquiries made to apply for new loans; and various geographic identifiers as detailed as the census block.[4]

[3]The randomness of the sample derives from including only individuals whose last two digits of their social security numbers belong to a pre-specified set of five numbers. The last four digits of the social security number are assigned sequentially to new applicants in chronological order as applications are processed and are thus as good as randomly assigned. Such randomness is a key feature of the research design of, for example, Johnson et al. (2006) and Gross et al. (2014).

[4]For more information about the CCP, see Lee and van der Klaauw (2010).

2.2 Background on Credit Scoring

Like most other credit scores, the Equifax Risk Score is designed to predict the likelihood of severe delinquency over the next 24 months but is estimated using an algorithm different from FICO's and ranges from 280 to 850, with a higher score indicating higher credit quality (lower default probability). In constructing credit scores, credit reporting agencies' algorithms consider an individual's debt payment history (including foreclosure and bankruptcy) as the primary determinant. Other factors, such as levels of indebtedness, age of the credit history, credit limit utilization, and new credit inquiries are also important inputs (Chatterjee et al., 2022).

Importantly, essentially all credit score algorithms exclude information on race, gender, age, income, assets, employment history, and occupation in estimating an individual's credit score (Beer et al., 2018). Existing research is inconclusive as to the extent to which income is correlated with credit scores. On the one hand, Beer et al. (2018) find a rather modest correlation of 0.3 to 0.4 in a random sample of consumers whose income and credit scores are both observed. Moreover, conditional on a parsimonious set of credit history variables, household income has essentially no additional explanatory power on credit scores. On the other hand, Albanesi et al. (2020) argue for a stronger correlation between the two and that income is the second most important predictor, following age, of credit scores.

2.3 Other Data Sources

One limitation of the Equifax CCP data is that the dataset does not provide any demographic and socioeconomic information besides age, partly due to federal laws prohibiting the collection and use of information on race, ethnicity, national origin, sex, and marital status in loan underwriting and in developing credit score models. We therefore augment the Equifax CCP data using the 2000 Decennial U.S. Census. Specifically, we use the block group-level estimates as proxies for those of the individuals who live in the same block groups. For example, we use the white and college-educated population share of a block group as proxies for individual race and educational attainment, and the block group median income for individual income. We assess the statistical merit of this approach in the econometric framework section. In addition, we use the Panel Study of Income Dynamics (PSID) and the Social Capital Community Survey (SCCS) to supplement and corroborate our main analysis.

3 Algorithm of Identifying Relationship Formation and Dissolution

3.1 An Address-Based Algorithm

Taking advantage of the large sample size in our data (roughly 12 million consumers in a typical quarter), we design a novel algorithm to identify committed relationships, when they form, and when they dissolve. The general idea of the algorithm is to first find pairs of individuals between 20 and 55 years old who started to share the same address in quarter Q, then screen the pairs and keep only those who continued to live together in the next four quarters but had not lived together in the eight quarters before Q.[5] In addition, we remove the pairs who shared the address with other individuals to filter out potential roommates and parents-children relationships.[6] Applying this algorithm to the primary-sample individuals yields a sample of near 50,000 committed relationships. For individuals in this sample, we can track them both before and after the relationship formation. In one of the robustness analyses, we also apply a modified version of this algorithm to include all consumers in the primary and non-primary CCP samples. This approach yields a much larger sample of over two million relationship, though not all individuals in this sample can be followed consistently over time.[7]

Importantly, we note that our algorithm of identifying committed relationships cannot distinguish married couples from those in a cohabiting relationship. However, this distinction

[5] In a robustness check, we strengthen the restriction that couples had not shared the same address 16 quarters prior to Q.

[6] Dettling and Hsu (2018) use a similar algorithm to find the adult children who moved back to parents' home in the same data.

[7] Because we only observe non-primary sample individuals when they share the same address with primary sample individuals, we define that a primary and non-primary consumer formed a committed relationship in quarter Q if Q is the first time that the non-primary consumer appears in the data.

is not all that critical for our analysis because we are interested in the implications of credit scores and the associated match quality in a general swathe of committed relationships, not just the couples who are legally married. Indeed, the marriage rate has experienced a steady decline since the 1990s, whereas cohabitation has become an acceptable living arrangement in the society (Lundberg and Pollak, 2014; Wang and Parker, 2014). Moreover, many cohabiting relationships eventually evolve into marriages, further blurring the distinctions between the two types of committed relationships (Stevenson and Wolfers, 2007). Lastly, cohabiting couples also share many household economic and financial responsibilities in a way similar to married couples, making credit scores an important subject to study even in this context.

Our strategy of identifying the relationship dissolution (separation) follows a similar idea. Specifically, among the identified relationships, we consider a couple as having separated in quarter Q if the two individuals begin living at different addresses for at least five quarters starting from Q and are never observed to share the same address again. Note that only separations within the sample of identified couples are studied as the objective of our analysis is to relate the outcome of a relationship to the conditions prevailing at the start of the relationship.

3.2 Validating the Algorithm

Because this is the first effort of identifying committed relationship using an address-based algorithm, we run an array of analysis to assess its validity and performance.

3.2.1 Relationship formation rates

We first study the relationship formation rates derived from our algorithm. As shown in Table 1, the formation rate in the Equifax primary sample is 0.108 percent during our sample period for individuals between 20 and 55. The low formation rate is due to the baseline algorithm identifying only the relationships formed among individuals of the primary sample. To see this, let Ω represent the population, let ω be a 5 percent random subset of Ω, and P_Ω be the national relationship formation rate. Individual $i \in \omega$ then has a probability P_Ω of forming a relationship with an individual $j \in \Omega$ in a given year. However, the probability of forming a relationship with $j \in \omega$, P_ω, should follow

$$P_\omega \approx 0.05 \times P_\Omega \tag{1}$$

Applying this adjustment gives the implied population relationship formation rate of 2.2 percent (shown in column 2). Indeed, when we apply the algorithm to include the non-primary sample, we obtain a relationship formation rate of 2.3 percent (column 3), similar to the adjusted primary sample formation rate. The formation rates implied by our algorithm are somewhat higher than the marriage rate of 1.5 percent among individuals of the same age group estimated using the Vital Statistics and the U.S. Census data. The gap likely reflects our algorithm picking up cohabiting relationships. Consistent with the presence of cohabiting couples, we note that the dissolution rate, reported in the lower panel, in the baseline sample is fairly high. About 15 percent of these couples separate by the end of the relationship's second year, and 15 percent of the couples whose relationship survive the first two years separate by the end of the fourth year.

For lifecycle trends, the relationship formation rate declines with age, with the rate among individuals aged between 20 and 35 being above 2.6 percent while the rate among individuals aged between 46 and 55 below 1.5 percent. We do not have administrative data to estimate marriage rates by age. However, the decline in relationship formation rate is broadly consistent with the estimates of marriage rate derived using household survey data. For example, the marriage rates estimated using the PSID data are about 1.5 percentage points higher among the subsample aged between 20 and 35 than for the subsample between 45 and 55, a difference similar to that in our estimates.

3.2.2 Do they look like couples?

We now examine whether individuals involved in the identified relationships demonstrate similarities regarding their demographic and socioeconomic characteristics in a way consistent with the couples in the household survey data. As the CCP data have only age information, we use census block group level statistics from the 2000 Census as proxies for individual characteristics.

As shown in column 1 of Table 2, the average age of the individuals at the time of relationship formation is about 37 years, the average age differential is 3.6 years, and the within-couple age correlation is 0.85. These statistics are broadly consistent with estimates from nationally representative surveys in which marital status is observed. For example, in the PSID data, the average age at marriage is slightly younger at about 34 years (column 2). Moreover, the age differentials and within-couple age correlation are remarkably similar between the CCP and the PSID estimates and are in line with previous estimates (Watson et al., 2004).

The correlations in partners' demographics are shown in the bottom panel of the table. The estimated within-couple racial correlation is above 0.6, the college degree correlation is slightly below 0.5, and the income correlation is 0.35. The magnitudes of the estimated correlations are consistent with previous studies, as is the rank ordering of their size (see, for example, respectively, Weiss and Willis (1997); Garfinkel et al. (2002); Blackwell and Lichter (2004)). Furthermore, the racial and income correlation coefficients are similar to those estimated using the 1999-2009 PSID sample of newly married couples. While the PSID within-couple college degree correlation appears to be somewhat lower than the CCP estimates, the latter are very similar to the correlation estimated using other nationally representative household surveys, such as the Survey of Consumer Finances (correlation coefficient = 0.52) and the Consumer Expenditure Survey (correlation coefficient = 0.47).[8]

Finally, we construct a placebo sample of randomly matched individuals in the primary sample that adhere to the age and age-difference restrictions imposed on the identified couples. Accordingly, as shown in column 3, the age-related statistics are very similar to those in column 1. However, we do not observe any correlations regarding race, educational attainment, and income within the placebo sample, a reassuring contrast for the validity of our algorithm.

[8] The PSID data demonstrate a persistent decline in the within-couple college degree correlation over time, making the PSID statistics somewhat lower than those derived from other national surveys.

4 Econometric Framework and Omitted Variable Concerns

Our analysis aims to address three questions: First, as a preliminary step to introduce credit score's relevance in the context of committed relationships, we ask whether credit scores predict relationship formation. Second, what is the match quality in couples' credit scores at the onset of the relationship, above and beyond other between-partner similarities? Third, do couples' credit scores (levels and match quality) predict relationship outcomes? Our empirical analysis proceeds in three parts, accordingly.

4.1 Relationship Formation

First, in the cross-section of single individuals, we analyze the likelihood of forming a committed relationship in a given year by estimating various specifications of the following model:

$$R_y^i = \alpha + \beta Score_y^i + \gamma Age_y^i + \theta Z^i + S^i + Year_y + u_y^i. \tag{2}$$

Here, R_y^i is a binary variable for whether individual i, who is single in the fourth quarter of year y, forms a committed relationship, as defined in Section 3.1, by the fourth quarter of year $y + 1$.[9] $Score_y$ is the level of the credit score or a vector of credit score bins for individual i, depending on the specification, in year y, when the individual remains single. Age_y is a cubic polynomial of the individual's age in year y. Z is a vector of proxies for individual socioeconomic characteristics that include the log of median income, share of college educated, and share of white residents of the census block group where individual i is observed to have resided the longest time prior to year y. S is a vector of state or county fixed effects to control for regional variations in how people sort into committed relationships as well as geographic patterns in credit scores.[10] $Year$ is a vector of year fixed effects to control for macroeconomic conditions and aggregate trends in relationship formation and credit scores. In some specifications, we

[9]Such relationships include those with nonprimary-sample individuals.

[10]For instance, on average, credit scores tend to be lower in the Southeast where marriage markets may also operate differently, as evidenced by the younger average age at first marriage.

allow for interacted fixed effects of $S \times Year$ to account for regional variations in these trends. The parameter of interest is β, and our conjecture is $\beta > 0$ as singles with higher credit scores tend to have greater access to credit and better consumption smoothing capabilities, which in turn makes them more attractive to a larger set of partners in the relationship market.

4.2 Match Quality upon Relationship Formation

For the second question, among the identified committed relationships, we measure match quality at the time of relationship formation and the extent to which we see positive sorting by credit scores. We also estimate the joint distribution of credit scores and compare the distribution with what would arise if individuals in the sample are randomly matched. We also analyze how couples' match quality evolves over time and contrast such dynamics between couples who stay together and those who separate.

4.3 Relationship Dissolution

Third, we estimate the following model to study how *initial* average credit scores and match quality are associated with the likelihood of separating and to explore the potential channels:

$$\begin{aligned} D^i_{q_1,q_2} \;=\;& \alpha + \beta_a \overline{Score}^i_{q_0} + \beta_g Score\ Gap^i_{q_0} + \gamma Age^i_{q_1} + \gamma_g\ Age^{gap,i} + \theta_g Z^{gap,i} + \delta Div^i \\ & + \nu CM^i_{q_1,q_2} + S^i + Year_{q_1} + u^i_{q_1,q_2}. \end{aligned} \tag{3}$$

Here, $D^i_{q_1,q_2}$ is a binary variable that indicates whether the couple, i, conditional on remaining in the relationship as of quarter q_1, separates by quarter q_2. Specifically, denoting q_0 as the quarter of relationship formation, we examine the prospects of relationship dissolution within three windows: $(q_1, q_2) \in \{(q_0 + 5, q_0 + 8), (q_0 + 9, q_0 + 16), (q_0 + 17, q_0 + 24)\}$. $\overline{Score}_{q_0}$ is the average credit score of the two partners measured at the time the committed relationship forms, and $Score\ Gap_{q_0}$ is the credit score difference observed at that time.

Among couples comprising two primary sample individuals, the longitudinal structure of the Equifax data allows us to follow both partners before and after the formation of their relationship. This feature offers critical advantages to circumvent some important endogeneity concerns in estimating the model. In particular, as two partners often share financial responsibilities and decisions, along with influencing each other on their attitudes and approaches of managing finances, their credit scores tend to converge during the course of their relationship. However, the credit scores of couples who separate do not converge, a pattern that we will explore in detail in the next section. Thus, the credit score match quality observed during a relationship can be correlated with the prospects of the relationship itself. In contrast, using *initial* credit score match quality observed at the time of relationship formation and relating this measure to *subsequent* relationship dissolution largely circumvents this endogeneity concern.[11]

Age and Age^{gap} denote age polynomials of both partners and their age difference, respectively. In addition, we include Z^{gap}, a vector of gaps in partners' socioeconomic and demographic proxies, to help us understand how such differences contribute to relationship dissolution. This inclusion also helps to control for factors that may confound the relationship between dissolution and credit scores. Lastly, we include the local divorce rate, Div, observed in the 2000 decennial Census to control for the potential social stigma of divorce or separation. For couple i, Div corresponds to the 2000 census block group-level divorce rate in the couple's neighborhood where the couple resides as of q_1 so that separations in our sample do not feed back into the control variable.

To study the channels by which initial credit score conditions influence future relationship outcomes, we include CM_{q_1,q_2}, a vector of credit market outcomes realized between q_1 and q_2, in the model. CM includes both measures of credit use, such as joint account ownership and new borrowing, and indicators of financial distress, such as bankruptcy, foreclosure, and other derogatory records. Indeed, mechanically credit scores are designed to predict such credit market experiences, which themselves have an effect on relationship outcomes. The coefficients of $\overline{Score}$ and $Score\ Gap$, therefore, speaks to our hypothesis that credit scores may affect relationship outcomes beyond credit market channels. Lastly, S^i and $Year_{q_1}$ are state of the

[11] In one of the robustness analyses, we use the credit score gaps observed eight quarters prior to the relationship formation to further isolate the potential comovements.

couple's residence prior to period (q_1, q_2) and calendar year fixed effects.

4.4 Control Variable Proxies and the Omitted Variable Biases

We leverage the detailed address information and use the census block group-level statistics of the 2000 Decennial U.S. Census as a proxy for individual demographic and income variables. We use block group-level proxies because credit bureau data do not have information on individual demographic characteristics and income. These variables may be correlated with credit scores and have an effect on relationship outcomes, implying potential omitted variable-biases in our estimates. Because a census block group is a rather small geography (typically with a population of approximately 1,000 people), we believe that such proxies can capture meaningful covariation between credit scores and demographic characteristics.

To shed light on the statistical merit of this approach, we estimate a stylized Mincer model using a PSID sample of household heads.[12] Assuming that we are interested in the white-nonwhite earnings gap, the model below also includes controls of age and educational attainments.

$$Log(Labor\ Income) = \alpha + \beta Age + \gamma White + \psi_1 Below\ HS + \psi_2 College + Year + \varepsilon, \qquad (4)$$

where Age is a third-order polynomial in household head age, $White$ is a dummy equal to 1 if head is white and zero otherwise, $Below\ HS$ and $College$ are head educational attainment dummies, with high school graduates being the omitted group.

The results are shown in column 1 of table 3, with a $White$ estimated coefficient of 0.256. Now, suppose that the education data were not included in the sample. The coefficients estimated from a model without controlling for educational attainments are in column 2. The $White$ coefficient becomes 0.341, 33 percent larger than in column 1, and the age polynomial coefficients are also amplified by a sizeable margin. The changes of the estimated coefficients highlight the omitted variable-bias. If, instead, we estimate the model using census-tract level shares of below-high-school and college-educated population from the Decennial U.S. Census

[12]Specifically, we use the PSID data from 1999–2009 and include household heads whose ages were between 20 and 65, currently working for someone else, and with an annual labor income greater than $5,000.

statistics as proxies of individual levels of education, the estimated coefficients in column 3 become considerably closer to the baseline results in column 1. In particular, the $White$ coefficient is 15 percent off, more than halving the margin of bias.[13] Moreover, the R^2 of the model rebounded from 0.168 to 0.251, only 8 percent below the baseline R^2 of 0.273.

Does the geography level of proxy matter for the performance of accounting for omitted variable bias? If we use the county-level, rather than the census tract-level, population shares, the coefficients shown in column 4 are very similar to those estimated without the approximated controls. On balance, our stylized exercises suggest that the proposed proxies do a decent job reducing the omitted variable bias, particularly when constructed at a granular geography (e.g. census tracts). We note that the proxies used in our main analysis are at the block group level, the population of which is typically a fraction of that of a census tract. Therefore, we expect the statistical performance of these proxies to be even better than those at the tract level with respect to addressing omitted variable bias.

Finally, we evaluate the quantitative magnitude of the omitted variable bias in the context of relationship dissolution. We use again the PSID data to estimate the following stylized, linear probability model that compares the divorce hazards of white-white couples against those of the couples with other racial compositions.

$$Divorce = \alpha + \beta Age + \gamma WW + \psi^h Edu^h + \psi^w Edu^w + \pi^h LogInc^h + \pi^w LogInc^w + Year + \varepsilon, \quad (5)$$

where $Divorce$ is a dummy equal to one if the couple divorced within a two-year period, Age represents third-order polynomials of both head and spouse ages, WW is equal to one for white-white couples and zero otherwise. Edu^h and Edu^w are vectors of head and wife educational attainments, and $LogInc^h$ and $LogInc^w$ are head and wife log labor earnings.[14]

As shown in table 4, without including any controls, the WW couples are estimated to

[13] The census tract-level location data are not included in the PSID public datasets. Researchers can apply for access to the restricted data.

[14] For individuals with zero labor earnings, we set these variables to zero. We also include dummies that indicate zero labor earnings. We use a sample of married couples in the PSID data from 1999 to 2009 with head age between 20 and 50.

have a 5 percentage points lower likelihood of divorcing in a two-year period than other couples (column 1). Adding head and spouse age polynomial leaves the estimated WW coefficient little changed (column 2). Further adding education controls dials the estimated γ down to 0.041. Finally, adding head and wife labor earning controls yields an estimated $\gamma = 0.039$ that remains statistically significant. On balance, this stylized exercise indicates that age, education, and income controls, while boosting the R^2 to some extent, do not qualitatively alter the estimated racial gap of divorce hazard. This analysis lends additional reassurance that our baseline analysis is unlikely qualitatively biased as a result of not having individual demographic and income controls.

5 Results

We have raised three empirical questions—do credit scores predict relationship formation? What is the match quality in couples' credit scores at the time of forming a relationship? Do couples' initial levels and match quality of credit scores predict relationship outcomes? This section presents the results estimated following the econometric approach introduced above.

5.1 Credit Scores and Relationship Formation

To begin, Table 5 presents the estimates of Equation (2), with standard errors clustered at the individual level. Here, we estimate both a linear effect of credit scores (normalized by the sample standard deviation) on the relationship formation likelihood during the subsequent year and a model of prime and nearprime credit score dummies (those with subprime credit scores being the omitted group).[15] Estimated marginal effects of our logit models suggest that individuals with higher credit scores are more likely to form committed relationships—a result that is evident across different specifications.[16] We estimate that a one standard-deviation higher credit score, about 100 points, corresponds to a 1.0 to 1.5 percentage points higher likelihood of entering a

[15]Prime refers to RiskScore above 720, nearprime between 620 and 719, and subprime below 620.

[16]In this and the following tables, unless otherwise noted, marginal effects of continuous variables are estimated using the dy/dx option in Stata.

committed relationship in the following year, which is about 10 percent of the mean likelihood (shown at the bottom of the table) in our sample. In the alternative specification, compared with singles with subprime credit scores, those with nearprime scores are between 1.2 to 2.0 percentage points more likely to form relationships during the next year, whereas those with prime scores are 2.3 to 3.6 percentage points more likely to form relationships. We also replace the nearprime and prime dummies with an array of 50-point score bins. The marginal effects estimated, shown in Figure 1, reveals a Λ-shaped profile, with the credit score effect peaking around 750.

These results are robust to a number of alternative specifications. In columns 1-2, only a polynomial of the individual's age is controlled for; in columns 3-4, year fixed effects and individual characteristics proxies are added to account for individuals sorting themselves into relationships by age, race, education, and earnings potential; state fixed effects are included in columns 5-6. To further control for the possible variation in relationship formation prospects associated with time and location, we include state $\times$ year, county, and county $\times$ year fixed effects, respectively, in linear probability models (reported in columns 7-9).[17] The estimated coefficients are about 0.01 across columns 7-9, which are all statistically significant. Compared with the mean of the dependent variable 0.125 (reported in the memo line at the bottom of the table), the linear model estimates are stable across various specifications and are broadly consistent with the marginal effect estimates reported in columns 1-6.

Finally, the estimated coefficients of individual characteristics proxies suggest that white singles with higher income tend to have a higher likelihood of forming a relationship. However, the college educated appear to have a lower formation rate. We also note that we find no evidence of larger-than-average changes of credit scores during the quarters leading up to the relationship formation, ruling out a widespread "window-dressing" of credit scores prior to forming a relationship.[18]

[17] We use linear probability model to accommodate the large numbers of fixed effects.

[18] Specifically, among individuals between 20 and 55 who entered relationship in Q, during the eight quarters that led up to Q credit scores on average increase 5 points, comparing with a 7-point average increase within

5.2 Credit Score Match Quality upon Relationship Formation

In the upper panel of Table 6, the average credit score of the couples is about 660 at the time of relationship formation. This is somewhat lower than the whole-sample average of 680 and likely reflects the younger-than-average age of those forming committed relationships. Regarding individuals' sorting into relationships, we find a large and positive within-couple credit score correlation of about 0.6 at the time of relationship formation. The magnitude of this correlation is consistent with the notion that access to credit and the potential individual traits reflected in credit scores are important components in determining couples' joint consumption and capability of weathering income shocks. Moreover, partners' credit score percentiles, rather than their levels, demonstrate the same pattern.[19]

One implication of such a strong, positive correlation is that the dispersion in credit scores at the couple level is similar to that at the individual level. To see this, note that the standard deviation of the credit score distribution across all individuals in our sample of identified couples is about 105. In a sample where individuals were randomly matched, the estimated standard deviation of the distribution of couples' average-credit score is about 75. However, in our sample of committed relationships, the standard deviation of couples' average credit score is much higher at 92. Because a couple's access to credit and its prices (interest rates and fees) are often determined by the lower score of the two individuals, we further examine the dispersion of the minimum score between two partners. Its standard deviation, at 105, is essentially identical to that across all individuals in our sample. Thus, the inequality in credit access and costs over individuals is largely preserved among the couples involving these individuals. The contingency matrix in the lower panel confirms that credit scores are positively correlated within a couple and that the positive correlation is much larger than what we would expect to observe if individuals were randomly matched to their partners (presented in parentheses in the contingency table).

eight quarters among all consumers of similar ages.

[19] The percentiles are calculated relative to the score distribution of all primary sample individuals in the quarter of relationship formation.

Importantly, the matching in credit scores is not perfect. The estimated average within-couple credit score differential remains sizeable (69 points), but substantially smaller than the estimated average differential between two randomly matched individuals (about 150 points). In addition, credit score differentials appear to vary with the average levels of credit scores of the couple. The percentile differential for the two inner quartiles of the average credit score distribution is on average about 20, comparing with an average 12 for the two outer quartiles, indicating that the match quality among low- and high-score couples is somewhat better than among middle-credit score couples. Finally, because credit scores can fluctuate from quarter to quarter, it is important to rule out that the within-couple score differential estimated at the time of relationship formation represents a temporary change. We therefore examine the average score taken over across quarters $Q-3$ and Q to filter out temporary variations. The results are qualitatively unchanged relative to those presented in Table 6.

To put the within-couple credit score correlation upon relationship formation in context, it is helpful to compare this correlation with the correlations of other characteristics of the couple. For example, as reported in the lower panel of Table 2, the correlations of being white, having college education, and individual labor income (of the first year of marriage) are 0.66, 0.31, and 0.38, respectively, for couples observed in the PSID data. Thus, the initial credit score correlation is about the same magnitude as the correlation in race but higher than that in education and income. Moreover, after projecting each spouse's credit score on an age polynomial and the same set of proxies of individual characteristics, we continue to find that the within-couple correlation coefficient of the regression residuals is 0.50. The residual correlation is only slightly lower than the unconditional correlation of credit scores, lending support to the notion that the within-couple credit score correlation is unlikely to be primarily driven by socioeconomic and demographic similarities. Rather, the results are consistent with a hypothesis that credit scores reflect some individual traits that were taken into account in partner selection—a subject we will explore further in the next section.

5.3 Match Quality Dynamics

We now illustrate how match quality in credit scores evolves during the course of the relationship, shedding some light on the interactions between credit score match quality and the durability of the relationship itself. To begin, the upper panel of Figure 2 presents the dynamics of each partner's credit score during the first 16 quarters of the relationship (for relationships that last at least this long), with the blue (orange) curve tracking the average score of the individuals with higher (lower) scores when the relationship begins. As the figure shows, the difference in credit scores narrows appreciably over the first 16 quarters, and the convergence is largely reflecting an increase in the credit score of the partners with lower initial scores.[20] The pattern is similar and more pronounced among couples who remain together longer. As shown in the lower panel, the average credit score gap among couples who remain together for at least 32 quarters (8 years) narrows to only 10 points by the eighth year of the relationship.

This convergence in credit scores likely reflects the shared financial responsibilities and mutually influencing behaviors and attitudes between partners rather than spurious or other mechanical, lifecycle-related factors (such as aging). Indeed, the top panel of Figure 3 shows no convergence of credit scores after couples separate. Similarly, among the placebo couples we constructed by randomly matching individuals subject to the same age restrictions, we find no credit score convergence.

5.4 Initial Credit Score Conditions and Relationship Outcomes

In Table 7, we present the association between *initial* credit score conditions and subsequent relationship dissolution. In the table, both *Initial score gap* and *Initial score average* are normalized by their respective sample standard deviation. Similar to Table 5, we show various specifications of controls and fixed effects to ensure the robustness and consistency of our results. Columns 1-2 include no controls and no fixed effects, columns 3-4 include controls and

[20]Interestingly, it is not rare for the within-couple ranking of credit scores to switch over time. After the first four years, over one-third of the couples see a reversal in the score ranking.

year fixed effects, and columns 5-6 further include state fixed effects. Several salient features emerge in the results. First, couples with higher average credit scores are less likely to separate. Our marginal effect estimates indicate that a one standard deviation increase in couples' average credit scores (about 90 points) corresponds to a likelihood of separation 3.0–4.2 percentage points lower during the second year of a relationship. The associated reduction of separation likelihood is over 5 percentage points during the third and fourth year and about 3 percentage points during the 5th and 6th year (shown in the second and third panel). In these estimates, the reductions are about 30 percent of the mean separation probability.

Second, initial match quality, defined as the gap between two partners' credit scores at the time of relationship formation, Q, is highly predictive of relationship separation, even controlling for the average couple credit score. The marginal effects show that a one standard deviation increase in the score differentials (about 66 points) corresponds to a likelihood of separation during the second year of the relationship that is 3 percentage points higher. The corresponding marginal effects are about 2.5 percentage points higher for the third or fourth year, and about 1 percentage point higher during the fifth or sixth year. The estimates are largely unchanged when controlling for couples' average credit scores using more flexible 50-point bins (columns 2, 4, and 6). Moreover, estimating linear probability models that include state×year or county fixed effect yields largely similar results (columns 7-8). Finally, we note that the estimated coefficients of the control variables (not shown) are consistent with the notion that worse match quality on socioeconomic and demographic characteristics tends to predict less stable relationships. For example, relationships with greater age and educational attainment differentials and relationships involving different races tend to have greater chances of separation. In addition, couples residing in communities with a higher historical share of divorced population are more likely to separate. Notably, the results presented in Table 7 highlight one aspect in which relationship outcomes are predictable with ex ante information, complementing the view that separations are largely driven by shocks during the course of relationship (Becker et al., 1977; Charles and Stephens, 2004). The estimated association between initial credit score conditions and relationship outcomes are robust to a wide range of alternative econometric, sample selection, and relationship identification algorithm specifications, which we discuss in detail in Appendix A.

Note that estimating a linear version of Equation (3) without initial credit score levels and gaps yields an R^2 of 0.047, whereas adding credit score levels and gaps boosts the R^2 by over 50 percent to 0.072. The increase in R^2 suggests credit scores bear additional explanatory power of relationship outcome beyond income and demographic proxies. Put the comparison in a different perspective, estimating the model using only initial score conditions renders an R^2 of 0.037, comparable with that of the model with only spousal age polynomials and age differential (0.043). In addition, the univariate R^2 of $\overline{Score}_{q0}$ and $Score\ Gap_{q0}$ are about the same (0.022 and 0.025 respectively), suggesting that the initial credit score levels and match quality have similar effects on relationship durability, which we will discuss below.

5.5 Level Versus Gap

Relationships with lower average scores or large score gaps are more likely to separate. But would a couple with similarly low scores be more or less likely to separate than a couple with more different scores but a higher average? To illustrate the relative effects of initial score levels and gaps, we put couples in six relationship type bins based on two spouses' credit scores—(Subprime, Subprime), (Subprime, Nearprime), (Subprime, Prime), (Nearprime, Nearprime), (Nearprime, Prime), (Prime, Prime). We then estimate the following variation of Equation (3):

$$\begin{aligned} D^i_{q_1,q_2} &= \alpha + \beta\, Type^i + \gamma Age^i_{q_1} + \gamma_g\ Age^{gap,i} + \theta_g Z^{gap,i} + \delta Div^i \\ &\quad + \nu CM^i_{q_1,q_2} + S^i + Year_{q_1} + u^i_{q_1,q_2}, \end{aligned} \tag{6}$$

with the (Nearprime, Nearprime)-type couples being the omitted group. The marginal effects implied by estimated β are illustrated in Figure 4. Focusing on the left panel, the estimated marginal effects indicate that the (Subprime,Prime) relationships have the highest early separation hazard—over 10 percentage points (60 percent of the mean) higher than the omitted group and also higher than the (Subprime, Subprime) relationships, which have the lowest average scores. In addition, the (Nearprime, Prime) relationships have a 1.5 percentage points (about 10 percent of the mean) higher hazard than the (Nearprime, Nearprime) relationships. Estimates in the middle and right panels also indicate (Subprime,Prime) types having the highest separation hazards later during the relationship. In addition, the (Nearprime, Prime) types tend to have a higher hazards than the (Nearprime, Nearprime) omitted group during the third and fourth years of the relationship. On balance, the results suggest that the match quality effect can dominate the level effect within certain ranges of credit scores and during certain stints of the relationship.

6 Exploring Potential Channels

We now explore the channels through which credit score levels and match quality may affect relationship outcomes. We consider channels related to both economic factors (access to credit and credit-related distress) and factors related to personal traits. The existing literature generally assigns stronger attribution to unfavorable economic outcomes during the relationship—as opposed to partners' individual characteristics or their match quality—in explaining relationship dissolution (e.g. Lyngstad and Jalovaara (2010)). Our results suggest that pre-relationship spousal characteristics also play a significant role in accounting for relationship outcomes.

6.1 The Credit Market Channel

Couples with higher, better matched credit scores tend to have better access to credit and can borrow at a lower interest rate. These couples are therefore in a stronger position to support their joint consumption, weather income shocks, and manage financial risks, which may in turn lead to more stable relationships.

6.1.1 Credit score conditions predict credit usage and financial distress

To test this hypothesis, we first show the link between the couples' initial credit score conditions and their subsequent credit usage and credit-related distress. Specifically, we estimate a variation of Equation (3) with the dependent variable being replaced with several credit market outcomes, namely opening joint accounts, acquiring new debt, experiencing financial distress. In addition to the couples' average credit scores at relationship formation, we also control for county-level changes in unemployment and house prices to better isolate local economic conditions' effects. The parameter of interest—marginal effects implied by the coefficients of credit score match quality—are reported in Table 8.

As shown in the upper panel, couples with worse match quality in credit scores are less likely to open a joint mortgage, auto, or credit card account in the first two or four years of the relationship. For example, a one standard deviation increase in the initial credit score differential is associated with a 4 percentage points-lower (35 percent of the mean) likelihood of having a joint mortgage in the first two years (column 1). The estimated marginal effects are around 3.0 and 1.1 percentage points, respectively, for opening joint auto loans and credit card accounts (columns 2-3). Looking over a four-year horizon, the estimated marginal effects imply a qualitatively similar relationship between match quality and the use of joint accounts. Turning to acquiring new credit, as shown in the middle panel, greater initial credit score gaps also predict a lower likelihood of borrowing new mortgage and auto loans (regardless whether the debt were borrowed jointly).[21] Moreover, we find that better matched couples are less likely to experience financial distress, especially in the first two years of the relationship. The results shown in the lower panel indicate that a one standard deviation increase in initial credit score gaps is associated with a 0.4, 0.1, and 1.2 percentage points higher likelihood of personal bankruptcy, home foreclosure, and derogatory records, respectively, which account for 17, 8, and 13 percent of the mean of dependent variables.

[21] Because of the revolving nature of credit cards, we focus on new debt borrowing only for mortgage and auto loans.

6.1.2 These credit events predict relationship dissolution

We then correlate relationship dissolution during the third and fourth year with credit market events of the first two years (upper panel of Table 9) and dissolution during the fifth and sixth year with credit market events of the first four years (lower panel), with the same controls as in column 5 of Table 7. The results indicate that observed joint account ownership and new borrowing are both strong predictors of lasting relationships. For example, the likelihood of separation during the subsequent two years among the couples who opened joint credit accounts during the first two years of their relationship are lower than a quarter of those who did not open any joint accounts (column 1). To a lesser degree, borrowing new mortgage or auto debt also predicts lower chances of separation (column 3). In contrast, as expected, financial distress, on balance, predicts a higher chance of separation. For example, shown in column 5, couples filing for personal bankruptcy in the first two years of the relationship are 5.4 percentage points more likely to separate over the next two years than other comparable couples (about 35 percent of the mean of dependent variable).

6.1.3 Beyond the Credit Market Channel

We now estimate an augmented Equation (3) with additional controls of observed credit use and credit-related distress. Note that the credit score related variables are measured ex ante at the onset of the relationship, while credit use and distress are measured during the course of the relationship. If credit scores predict relationship outcomes only through the credit market channels discussed above, they should not have a sizeable, significant coefficient in such a model. The results, reported in columns 2, 4, and 6 of Table 9 demonstrate two notable features. First, the estimated marginal effects associated with credit score match quality are smaller than those in Table 7, suggesting that some of its predictive power indeed plays through predicting credit usage and financial distress. Second, and importantly, the results consistently show that initial credit score gaps predict future separations beyond the observed credit usage and financial distress. Finally, we estimate a "kitchen sink" style model that include all observed credit usage and financial distress as controls, and, as shown in column 7, credit score match quality continues to have sizeable, significant predictive power.

In an alternative specification, we first estimate linear probability models of separation dummies on credit usage and financial distress indicators. The residuals of these regressions pick up the component of dissolution likelihood not accounted for by credit market channels. We then correlate the residuals of these regressions with credit score gaps. The results (not shown) are similar to those in Table 9 and continue pointing to match quality's predictive power beyond the observed credit use and financial distress.

6.2 Interpretation—A Behavioral Traits Hypothesis

The results above point to credit scores being a potent predictor of relationship dissolution beyond their links with the credit events experienced by the couples. We ask the question what other signals may be reflected in individuals' credit scores that help predict their relationship outcomes. We conjecture that credit scores reveal one's underlying behavioral traits that we refer to as "trustworthiness." Our hypothesis shares a similar spirit with the existing survey-based results that conscientiousness, one of the big-five personal traits, has a significant influence on one's relationship (Boertien et al., 2015).

The idea can be illustrated in the following conceptual framework. Broadly speaking, a person's default probability is affected by her willingness to repay debts, which we denote with ζ, and a shock to her ability to repay, η.

$$default = f(\zeta) + \eta, \tag{7}$$

where ζ is, in turn, closely related to an individual's underlying trustworthiness. Consistent with such a framework, Chatterjee et al. (2022) noted that "individual's true propensity to repay—i.e., the individual's true type—is hidden from her creditors, and it is the presence of this persistent hidden information that makes an individual's history of actions relevant for lenders." Indeed, lenders have long recognized a borrower's general trustworthiness and character as important factors influencing her debt payment history and default probability. The credit reports garnered in the 1930s, in addition to debt repayment history and other financial data, collected information on borrowers' characteristics, reputation, habits, morals, and even illegal liquor traffic activities. For example, one of the credit reports prepared by the Retail Credit Company in 1934 included the following questions: "Does his record show he has been a steady and reliable man?" "Is his personal reputation as to character, honesty, and fair dealing good?" "Do you learn any illegal liquor traffic activities or domestic difficulties?"[22]

Modern credit reports no longer collect such information. However, to the extent that past defaults reflect willingness to repay debt, credit scores likely remaining informative regarding underlying trustworthiness. Because household relationships are intrinsically subject to fewer formal, contractual restraints and use more implicit contracts that can be difficult to enforce, one's trustworthiness are natural important predictors of successful relationships.

In addition, while adverse financial and economic shocks affect debt payment capability, thereby affecting consumers' credit scores to a significant extent, a substantial portion of credit score variations are not accounted for by such shocks and these residual components can predict relationship solutions. Specifically, we project the initial credit score of each partner on lagged county-level unemployment, delinquency, and personal bankruptcy rates, up to eight quarters, to decompose the initial credit score conditions into a predicted and a residual component that we include in the model of relationship dissolution. The results, reported in Table 10, show that, through the duration of the relationship, the residual component is consistently associated with greater likelihood of subsequent separations. By contrast, the predictable components do not have a consistent predictive power. Therefore, it is unlikely that the link between credit scores and future dissolutions solely reflects previous labor and credit market experiences occurred before they form the relationship.

So far, empirical analysis of establishing credit scores as an indicator for individual trust-

[22]The images of these credit reports are available from the authors upon request. See also Lauer (2017) for a detailed history of credit reporting.

worthiness remains at a very young phase. We report in Appendix B an array of suggestive evidence that demonstrates the similarity between credit scores and survey-based measures of trustworthiness. Both measures are capable of predicting relationship outcomes.

7 Conclusion

This paper brings credit scores to the front and center of the analysis on committed relationship formation and dissolution. Specifically, we establish three novel empirical patterns using a large proprietary dataset of the credit records of U.S. consumers over fifteen years. First, higher credit scores imply a greater likelihood of forming a relationship within the next twelve months; second, partners' credit scores are highly correlated at the time of relationship formation, even conditioning on socioeconomic and demographic similarities; and third, those with lower average and worse match quality of credit scores at the onset of relationships are more likely to separate within the subsequent several years.

While such a predictability suggests a plausible causal effect, we underscore that our research design does not fully establish the causality. That said, the fact that initial credit score conditions predict relationship outcomes is a new and potentially significant contribution to the literature. In addition, our sample was constructed using a near 5 percent random sample of the U.S. adult population over a long period of time that includes over a large number of committed relationships. Accordingly, the statistical results are likely generalizable to the broad population.

Regarding the channels in which initial credit score conditions may influence subsequent relationship outcomes, we first show that initial credit score conditions predict future credit usage and credit-related financial distress, which in turn may affect relationship durability. However, such a credit-market channel does not account for the entire predictability, prompting us to contemplate a conjecture that credit scores reveal an individual's underlying behavioral traits (such as trustworthiness) that also has a bearing on relationship outcomes. We underscore that this conjecture, while intuitive and promising, needs much more empirical analysis to validate.

Indeed, future research is warranted to explore the link between credit scores and personal traits and their potential as an algorithm-based measure of trustworthiness derived from people's economic behaviors. Such research will push the boundaries of our understanding of individual and social trust. After all, as Putnam (1995) famously wrote, "since trust is so central to the theory of social capital, it would be desirable to have strong behavioral indicators of trends in social trust and misanthropy. I have discovered no such behavioral measures."

References

Albanesi, Stefania, Giacomo DeGiorgi, and Jaromir Nosal, "Growth and the Financial Crisis: A New Narrative," *Working Paper*, 2020.

Alesina, Alberto and Eliana La Ferrara, "Who Trusts Others?," *Journal of Public Economics*, 2002, *85*, 207–234.

Arpino, Bruno, Marco Le Moglie, and Mencarini Letizia, "What Tears Couples Apart: A Machine Learning Analysis of Union Dissolution in Germany," *Demography*, 2022, *59* (1), 161–186.

Arrow, Kenneth, "Gifts and Exchanges," *Philosophy and Public Affairs*, 1972, *1* (4), 343–362.

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune, "Marry for What? Caste and Mate Selection in Modern India," *American Economic Journal: Microeconomics*, 2013, *5* (2), 33–72.

Becker, Gary, Elizabeth Landes, and Robert Michael, "An Economic Analysis of Marital Instability," *Journal of Political Economy*, 1977, *85*, 1141–1187.

Beer, Rachael, Felicia Ionescu, and Geng Li, "Are Income and Credit Scores Highly Correlated," *FEDS Notes 2018-08-13*, 2018.

Blackwell, Debra and Daniel Lichter, "Homogamy among Dating, Cohabiting, and Married Couples," *The Sociological Quarterly*, 2004, *45* (4), 719–737.

Boertien, Diederik, Christian von Scheve, and Mona Park, "Can Personality Explain the Educational Gradient in Divorce? Evidence From a Nationally Representative Panel Survey," *Journal of Family Issues*, 2015, *38* (10), 1339–1362.

Bos, Marieke, Andres Liberman, and Emily Breza, "The Labor Market Effects of Credit Information," *Review of Financial Studies*, 2018, *31* (6), 2005–2037.

Bricker, Jesse and Geng Li, "Credit Scores, Social Capital, and Stock Market Participation," *FEDS Working Paper No. 2017-008*, 2017.

_ and _ , "Your Friends, Your Credit: Social Capital Measures Derived from Social Media and the Credit Market," *FEDS Working Paper No. 2023-048*, 2023.

Brown, Jeffrey, Zoran Ivkovic, Paul Smith, and Scott Weisbenner, "Neighbors Matter: Causal Community Effects and Stock Market Participation," *Journal of Finance*, 2008, *43* (3), 1509–1531.

Charles, Kerwin and Melvin Stephens, "Job Displacement, Disability, and Divorce," *Journal of Labor Economics*, 2004, *22* (2), 489–522.

_ , Erik Hurst, and Alexandra Killewald, "Marital Sorting and Parental Wealth," *Demography*, 2013, *50* (1), 51–70.

Chatterjee, Satyajit, Dean Corbae, Kyle Dempsey, and José-Víctor Ríos-Rull, "A Quantitative Theory of the Credit Score," *Working Paper*, 2022.

Chen, Daphne, Dean Corbae, and Andy Glover, "Can Employer Credit Checks Create Poverty Traps? Theory and Policy Evaluation," *Unpublished manuscript*, 2013.

Chetty, Raj, Matthew Jackson, Theresa Kuchler, Johnnes Stroebel, Abigail Hiller, Sarah Oppenheimer, and Opportunity Insights Team, "Social Capital and Economic Mobility," *Opportunity Insights Non-Technical Research Summary*, 2022.

Chiappori, Pierre Andre, Sonia Oreffice, and Climent Quintana-Domeque, "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market," *Journal of Political Economy*, 2012, *120* (4), 659–695.

Dettling, Lisa and Joanne Hsu, "Returning to the Nest: Debt and Parental Coresidence among Young Adults," *Labour Economics*, 2018, *54*, 225–236.

Ettin, Erika, "A Little Number Goes a Long Way," *Philadelphia Inquirer*, 2013.

Garfinkel, Irwin, Dana Glei, and Sara McLanahan, "Assortative Mating among Unmarried Parents: Implications for Ability to Pay Child Support," *Journal of Population Economics*, 2002, *15* (3), 417–432.

Glaeser, Edward, David Laibson, Jose Scheinkman, and Christine Soutter, "Measuring Trust," *Quarterly Journal of Economics*, 2000, *115* (3), 811–846.

Gottman, John Mordechai, *What Predicts Divorce? The Relationship between Marital Processes and Marital Outcomes*, Psychology Press, New York and London, 2014.

Gross, Tal, Matthew Notowidigdo, and Jialan Wang, "Liquidity Constraints and Consumer Bankruptcy: Evidence from Tax Rebates," *Review of Economic Statistics*, 2014.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "The Role of Social Capital in Financial Development," *American Economic Review*, 2004, *94* (3), 526–556.

Johnson, David, Jonathan Parker, and Nicholas Souleles, "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 2006, *95* (5), 1589–1610.

Karlan, Dean, "Using Experimental Economics to Measure Social Capital and Predict Financial Decisions," *American Economic Review*, 2005, *83* (5), 1688–1699.

Lauer, Josh, *Creditworthy: A History of Consumer Surveillance and Financial Identity in America*, Columbia University Press, 2017.

Lee, Donghoon and Wilbert van der Klaauw, "An Introduction to the FRBNY Consumer Credit Panel," *Federal Reserve Bank of New York Staff Reports*, 2010.

Lundberg, Shelly and Robert Pollak, "Cohabitation and the Uneven Retreat from Marriage in the U.S., 1950-2010," in Leah Platt Boustan, Carola Frydman, and Robert A. Margo, eds., *Human Capital in History*, National Bureau of Economic Research, 2014.

Lyngstad, Torkild Hovde and Marika Jalovaara, "A Review of the Antecedents of Union Dissolution," *Demographic Research*, 2010, *23*, 257–292.

Mortelmans, Dimitris, ed., *Divorce in Europe: New Insights in Trends, Causes, and Consequences of Relationship Break-ups*, Springer Open, 2020.

Putnam, Robert, "Bowling Alone," *Journal of Democracy*, 1995, pp. 65–78.

Sapienza, Paola, Anna Toldra, and Luigi Zingales, "Understanding Trust," *The Economic Journal*, 2013, *123*, 1313–32.

Silver-Greenberg, Jessica, "Perfect 10? Never Mind That. Ask Her for Her Credit Score.," *New York Times*, 2012.

Solomon, Brittany and Joshua J. Jackson, "Why Do Personality Traits Predict Divorce? Multiple Pathways through Satisfaction," *Journal of Personality and Social Psychology*, 2014, *106* (6), 978–996.

Stevenson, Betsey and Justin Wolfers, "Marriage and Divorce: Changes and Their Driving Forces," *Journal of Economic Perspectives*, 2007, *21* (2), 27–52.

Wang, Wendy and Kim Parker, "Record Share of Americans Have Never Married," *Pew Research Report*, September 2014.

Watson, David, Eva Klohnen, Alex Casillas, Ericka Nus Simms, Jeffrey Haig, and Diane Berry, "Match Makers and Deal Breakers: Analyses of Assortative Mating in Newlywed Couples," *Journal of Personality*, 2004, *72* (5).

Weiss, Yoram, "The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And What Happens Upon Divorce?," in M.R. Rosenzweig and O. Stark, eds., *Handbook of Population and Family Economics*, Amsterdam: Elsevier Science B.V., 1997.

_ **and Robert Willis**, "Match Quality, New Information, and Marital Dissolution," *Journal of Labor Economics*, 1997, *15* (1), S293–S329.

Table 1: Annual Rates of Committed Relationship Formation and Separation

	Formation rates		
	Primary sample only		Whole sample
	Unadjusted (1)	Adjusted (2) = (1) × 20	(3)
Age 20-55	0.108%	2.16%	2.26%
Age 20-35	0.131%	2.62%	2.93%
Age 36-45	0.116%	2.32%	2.35%
Age 46-55	0.068%	1.36%	1.27%
Number of couples identified	49,363		2,070,117

	Separation rates	
	Primary sample only	Whole sample
During the 2nd year	15.1%	19.2%
During the 3rd and 4th year	14.9%	21.1%
During the 5th and 6th year	8.1%	12.4%

Note: For each age range, formation rates in column 1 are estimated as the ratios between the number of primary-sample individuals involved in committed relationships, as identified in the algorithm discussed in Section 2, and the total number of primary sample individuals. Column 2 presents the adjusted relationship formation rates, accounting for the fact that relationships formed between two primary sample individuals represent about 5 percent of the formation rates in the population. Column 3 presents the relationship formation rates estimated using the sample of couples identified when non-primary sample individuals are included. For each age range, these rates are estimated as the ratios between the number of primary-sample individuals involved in relationships identified using the algorithm allowing for non-primary sample individuals discussed in Section 2 and the total number of primary sample individuals. Separation rates are estimated as the ratio between the number of separated couples and the number of all relationships that survived through the first, second, and fourth years.

Table 2: Demographic Characteristics at the Time of Relationship Formation

	CCP data (1)	PSID data (2)	Placebo couples (3)
Individual level characteristics			
Average age	36.7	33.5	36.1
Age difference	3.6	3.8	3.7
Age correlation	0.85	0.86	0.82
Census block group level characteristics			
% White correlation	0.63	0.66	0.01
% College degree Correlation	0.48	0.31	-0.00
Median Income Correlation	0.35	0.38	0.02

Note: Column 1 presents the results corresponding to committed relationships among two primary sample individuals. Column 2 presents the results of the couples observed in the PSID data who are between the ages of 20 and 55. Column 3 presents the results of the placebo couples who are randomly matched in the CCP data subject to the same age and age differential restrictions. Age in the CCP data is calculated using the year-of-birth variable. Race, education, and median income variables are the census block group level statistics in the 2000 U.S. Census.

Table 3: Statistical Merit of the Control Proxies

	(1)	(2)	(3)	(4)
Individual Char.				
White	0.256***	0.341***	0.217***	0.332***
	(0.007)	(0.007)	(0.008)	(0.008)
Below HS	-0.302***			
	(0.010)			
College	0.384***			
	(0.008)			
Block-group proxies				
Share(Below HS)			-0.489***	
			(0.031)	
Share(College)			1.157***	
			(0.029)	
County proxies				
Share(Below HS)				-0.330***
				(0.050)
Share(College)				1.539***
				(0.047)
Controlling for				
Age polynomials	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
R^2	0.273	0.168	0.251	0.206
N	27,560	27,560	27,560	27,560

Note: Column 1 presents the estimates of regressing log labor income on observed racial and educational characteristics—a white dummy, a below high school dummy, and a dummy of college education. Column 2 presents the results excluding education attaintment dummies. Column 3 presents the results estimated using census block group-level proxies of educational attainments. Column 4 present the results estimated using county-level proxies. All models control for year fixed effects. *** denotes the estimate is statistically significant at the 99 percent level.

Table 4: An Assessment of the Omitted Variable Biases

	(1)	(2)	(3)	(4)
White-white couples	-0.051***	-0.050***	-0.041***	-0.039***
	(0.004)	(0.004)	(0.005)	(0.005)
Controlling for				
Age polynomials	No	Yes	Yes	Yes
Education	No	No	Yes	Yes
Labor income	No	No	No	Yes
Year F.E.	Yes	Yes	Yes	Yes
R^2	0.007	0.025	0.028	0.029
N	18,206	18,206	18,206	18,206

Note: Column 1 presents the estimates of a linear probability model, regressing a divorce dummy on a white-white couple dummy, controlling for only year fixed effects; column 2 adds head and wife age polynomials; column 3 adds head and wife educational attainments, and column 4 adds head and wife labor income as controls. *** denotes the estimate is statistically significant at the 99 percent level.

Table 5: Credit Score Levels and Spousal Relationship Formation

	Logit Report marginal effects (dy/dx)						OLS Report coefficients		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Credit score	0.015***		0.010***		0.010***		0.009***	0.010***	0.009***
(normalized with sample std.)	(0.000)		(0.000)		(0.000)		(0.000)	(0.000)	(0.000)
Nearprime		0.020***		0.012***		0.012***			
		(0.000)		(0.000)		(0.000)			
Prime		0.036***		0.023***		0.023***			
		(0.000)		(0.000)		(0.000)			
Log(median income)					0.019***	0.020***	0.019***	0.018***	0.018***
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share college					-0.005***	-0.005***	-0.005***	-0.005***	-0.005***
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share white					0.002***	0.002***	0.002***	0.001***	0.001***
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Control for									
Age polynomial	yes	yes	yes	yes	yes	yes	yes	yes	yes
Yearly FE	no	no	yes	yes	yes	yes	no	yes	no
State FE	no	no	no	no	yes	yes	no	no	no
Year × state FE	no	no	no	no	no	no	yes	no	no
County FE	no	no	no	no	no	no	no	yes	no
Year × county FE	no	no	no	no	no	no	no	no	yes
Pseudo R^2 and R^2	0.027	0.028	0.055	0.055	0.056	0.056	0.043	0.043	0.046
N	12,527,381	12,527,381	12,527,381	12,527,381	12,527,381	12,527,381	12,527,381	12,527,381	12,527,381
Mean of Dep. Var.									
Prob(relationship formation)					0.125				

Note. The table reports the marginal effects estimated from logistic regressions (dy/dx) in columns 1-6 and linear regression coefficients in columns 7-9. Standard errors are clustered at the individual level and reported in parentheses. Credit scores, log of median income, share of college educated, and share of white residents are normalized by respective sample standard deviations. *** denotes the estimate is statistically significant at the 99-percent level. Control variables of log median income, share of college educated, and share of white residents are census block group level statistics from the 2000 U.S. Census.

Table 6: Matching Quality at the Time of Household Formation

	Credit scores	
	Score levels	Score percentiles
	(1)	(2)
Mean credit score	657	41
Credit score standard deviations		
Cross-individual	104	26
Cross-couple (mean)	92	24
Cross-couple (minimum)	105	24
Within-couple correlation	0.59	0.63
Within-couple correlation (score residuals)	0.50	0.55
Mean within-couple score differentials	69	17

	Contingency table: observed versus randomly matched percentage of randomly matched are reported in parentheses		
	Prime	Near prime	Subprime
Prime	22.67 (11.73)	8.28 (10.30)	3.74 (12.66)
Near prime	7.58 (9.75)	12.56 (8.04)	8.68 (10.52)
Subprime	3.56 (12.34)	8.85 (10.83)	24.07 (13.32)
Column total	33.82	29.69	36.50

Note: The upper panel presents statistics estimated for the identified couples that involve two primary-sample individuals. Score percentiles are calculated with respect to the population credit score distribution of the quarter of relationship formation. Score residuals are the regression residuals of projecting credit scores on an age polynomial and proxies of individual characteristics of the same individuals. The lower panel presents the shares of couples by each spouse's credit score bin. Individuals with credit scores above 720 are labelled as prime, 620-719 as near prime, and below 620 as subprime. The numbers not in parentheses are observed shares, and the numbers in parentheses are the shares with individuals with different credit scores are matched randomly.

Table 7: Initial Credit Score Conditions and Subsequent Relationship Dissolutions

	Logit Report marginal effects (dy/dx)						OLS Report coefficients	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	The second year of relationship							
Initial score gap	0.036***	0.034***	0.031***	0.030***	0.031***	0.029***	0.036***	0.035***
(normalized by sample std.)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Initial score average	-0.042***		-0.030***		-0.031***		-0.027***	-0.029***
(normalized by sample std.)	(0.002)		(0.002)		(0.002)		(0.002)	(0.002)
Initial score average bins	No	Yes	No	Yes	No	Yes	No	No
R2	0.041	0.047	0.105	0.105	0.107	0.107	0.109	0.136
N	43,889	43,889	40,323	40,323	40,125	40,125	41,676	41,233
Mean of Dep. Var.								
Prob(separation)	0.151							
	The third and fourth year of relationship							
Initial score gap	0.031***	0.029***	0.026***	0.026***	0.025***	0.025***	0.031***	0.031***
(normalized by sample std.)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Initial score average	-0.057***		-0.050***		-0.051***		-0.052***	-0.531***
(normalized by sample std.)	(0.002)		(0.002)		(0.002)		(0.002)	(0.002)
Initial score average bins	No	Yes	No	Yes	No	Yes	No	No
R2	0.057	0.058	0.103	0.103	0.107	0.106	0.107	0.146
N	32,157	32,157	29,335	29,335	29,191	29,191	29,184	28,654
Mean of Dep. Var.								
Prob(separation)	0.149							
Controls								
Age polynomials	no	no	yes	yes	yes	yes	yes	yes
Age differentials	no	no	yes	yes	yes	yes	yes	yes
Current community char.	no	no	yes	yes	yes	yes	yes	yes
Community div. pop. share	no	no	yes	yes	yes	yes	yes	yes
Orig. community char. diff.	no	no	yes	yes	yes	yes	yes	yes
Yearly FE	no	no	yes	yes	yes	yes	no	yes
State FE	no	no	no	no	yes	yes	no	no
Yearly×state FE	no	no	no	no	no	no	yes	no
County FE	no	no	no	no	no	no	no	yes

Continued on next page

Table 7 – *Continued*

	Logit						OLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	The fifth and sixth year of relationship							
Initial score gap	0.011***	0.010***	0.010***	0.008***	0.009***	0.008***	0.011***	0.010***
(normalized by sample std.)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Initial score average	-0.030***		-0.030***		-0.031***		-0.033***	-0.356***
(normalized by sample std.)	(0.002)		(0.002)		(0.002)		(0.002)	(0.002)
Initial score average bins	No	Yes	No	Yes	No	Yes	No	No
R2	0.034	0.038	0.069	0.072	0.075	0.078	0.069	0.126
N	24,052	24,052	20,620	20,620	20,526	20,526	20,518	19,914
Mean of Dep. Var.								
Prob(separation)					0.081			
Controls								
Age polynomials	no	no	yes	yes	yes	yes	yes	yes
Age differentials	no	no	yes	yes	yes	yes	yes	yes
Current community char.	no	no	yes	yes	yes	yes	yes	yes
Community div. pop. share	no	no	yes	yes	yes	yes	yes	yes
Orig. community char. diff.	no	no	yes	yes	yes	yes	yes	yes
Yearly FE	no	no	yes	yes	yes	yes	no	yes
State FE	no	no	no	no	yes	yes	no	no
Yearly×state FE	no	no	no	no	no	no	yes	no
County FE	no	no	no	no	no	no	no	yes

Note. The table reports marginal effects (dy/dx) estimates in columns 1-6 and regression coefficients in columns 7-8. Standard errors are reported in parentheses. Credit scores are normalized by the sample standard deviation. *** denotes the estimate is statistically significant at the 99-percent level. For logistic regression, report pseudo R-squared. The estimations reported in columns 3-6 control for a third-order age polynomial of each spouse and the age differential. These estimations also control for differentials in socioeconomic characteristics, using differences of U.S. Census statistics of race, education, and income between the census block groups where each spouse lived prior to forming the relationship as proxies. These estimations also control for the share of divorced population where the couple lived as of the end of the first, second, and fourth year of their relationship, respectively, for the three sets of results.

Table 8: Initial Credit Score Differentials and Subsequent Credit Usage and Financial Distress

	First two years			First four years		
	(1)	(2)	(3)	(4)	(5)	(6)
	Open new joint loan accounts					
	Mortgage	Auto loan	Credit card	Mortgage	Auto loan	Credit card
Initial score gap	-0.040***	-0.030***	-0.011***	-0.050***	-0.022***	-0.009***
(normalized by sample std.)	(0.003)	(0.003)	(0.002)	(0.005)	(0.004)	(0.003)
Pseudo R2	0.121	0.052	0.078	0.129	0.047	0.067
N	20,426	22,953	22,201	13,134	15,588	14,610
Mean of Dep. Var.	0.113	0.096	0.038	0.204	0.147	0.059
	Borrow new debt					
	Mortgage	Auto loan	Credit card	Mortgage	Auto loan	Credit card
Initial score gap	-0.010**	-0.014***		-0.015**	-0.010*	
(normalized by sample std.)	(0.004)	(0.005)		(0.006)	(0.006)	
Pseudo R2	0.099	0.037		0.110	0.034	
N	13,217	12,182		8,492	8,697	
Mean of Dep. Var.	0.222	0.277		0.351	0.374	
	New financial distress					
	Bankruptcy	Foreclosure	Derogatory records	Bankruptcy	Foreclosure	Derogatory records
Initial score gap	0.004***	0.001	0.012***	0.000	0.001	0.002
(normalized by sample std.)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Pseudo R2	0.157	0.134	0.185	0.150	0.146	0.172
N	23,535	25,742	27,175	16,839	18,434	19,188
Mean of Dep. Var.	0.024	0.012	0.093	0.040	0.018	0.091

Note. The table reports marginal effect (dy/dx) estimates. Standard errors of estimated marginal effects are reported in parentheses. Credit scores are normalized by the sample standard deviation. *, **, and *** denotes the estimate is statistically significant at the 90, 95, and 99 percent level, respectively. The estimations control for the array of bins of couple-average credit score at the time of relationship formation. The estimations control for a third-order age polynomial of each spouse and the couple age differential; These estimations also control for differentials in socioeconomic characteristics, using differences of U.S. Census statistics of race, education, and income between the census block groups where each spouse lived prior to forming the relationship as proxies. Finally, the estimations control for yearly and state fixed effects.

Table 9: The Role Played by Use of Credit and Financial Distress on Separations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Separated during the third or fourth year of relationship						
Initial score gap		0.021***		0.025***		0.025***	0.021***
(normalized by sample std.)		(0.002)		(0.002)		(0.002)	(0.002)
Use of credit indicators							
Joint accounts	-0.192***	-0.159***					-0.158***
	(0.008)	(0.008)					(0.008)
Mortgage			-0.052***	-0.039***			-0.008
			(0.007)	(0.007)			(0.007)
Auto loans			-0.021***	-0.010			0.005
			(0.007)	(0.006)			(0.006)
Financial distress indicators							
Bankruptcy filing					0.054***	0.006	0.002
					(0.012)	(0.012)	(0.012)
Foreclosure					0.009	0.011	0.010
					(0.017)	(0.017)	(0.016)
Derog. records					0.075***	0.024***	0.021***
					(0.006)	(0.006)	(0.006)
Pseudo R2	0.099	0.127	0.073	0.108	0.077	0.107	0.127
N	30,228	29,191	30,228	29,191	30,228	29,191	29,191
Mean of Dep. Var.							
Prob(separation)				0.149			
	Separated during the fifth or sixth year of relationship						
Initial score gap		0.007***		0.008***		0.008***	0.007***
(normalized by sample std.)		(0.002)		(0.002)		(0.002)	(0.002)
Use of credit indicators							
Joint accounts	-0.070***	-0.057***					-0.052***
	(0.006)	(0.006)					(0.006)
Mortgage			-0.044***	-0.041***			-0.026***
			(0.006)	(0.006)			(0.006)
Auto loans			0.007	0.008			0.015***
			(0.005)	(0.005)			(0.005)
Financial distress indicators							
Bankruptcy filing					0.027***	0.010	0.005
					(0.009)	(0.009)	(0.009)
Foreclosure					0.019	0.019*	0.018
					(0.012)	(0.012)	(0.012)
Derog. records					0.042***	0.015**	0.013**
					(0.006)	(0.006)	(0.006)
Pseudo R2	0.065	0.088	0.056	0.082	0.057	0.079	0.090
	21,026	20,526	21,016	20,526	21,026	20,526	20,526
Mean of Dep. Var.							
Prob(separation)				0.081			

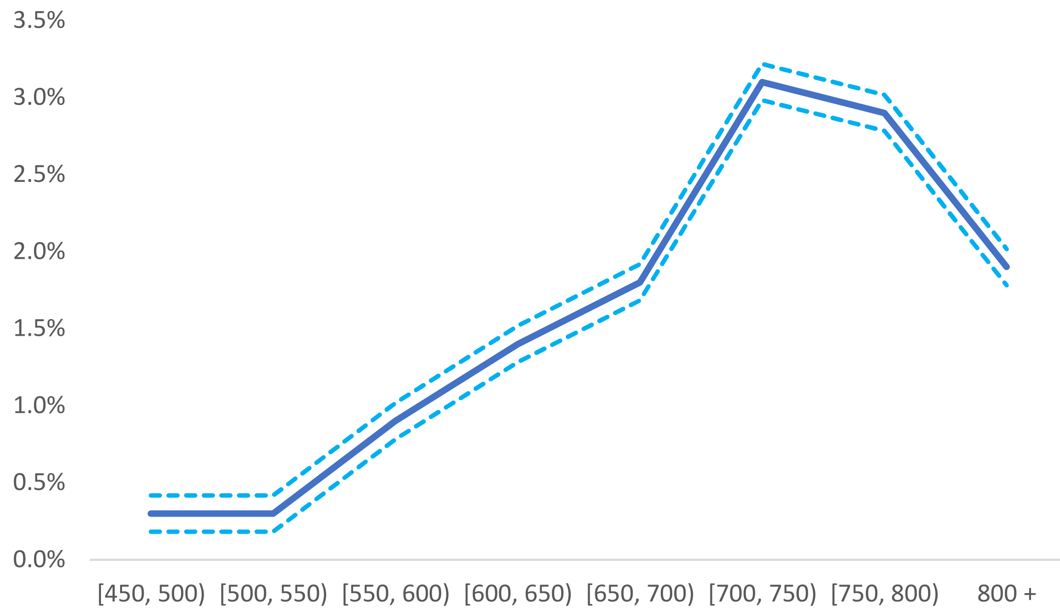
Note. The table reports marginal effect (dy/dx) estimates. Standard errors of estimated marginal effects are reported in parentheses. Credit scores are normalized by the sample standard deviation. *, **, and *** denotes the estimate is statistically significant at the 90, 95, and 99 percent level, respectively. The estimations control for the array of bins of couple-average credit score at the time of relationship formation. The estimations control for a third-order age polynomial of each spouse and the couple age differential; These estimations also control for differentials in socioeconomic characteristics, using differences of U.S. Census statistics of race, education, and income between the census block groups where each spouse lived prior to forming the relationship as proxies. Finally, the estimations control for yearly and state fixed effects.

Table 10: Predicted and Residual Score Levels and Gaps

	The second year				The third and fourth year				The fifth and sixth year			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Predicted score level	0.002				-0.008**				-0.001			
	(0.003)				(0.004)				(0.004)			
Residual score level		-0.038***				-0.057***				-0.031***		
		(0.002)				(0.002)				(0.002)		
Predicted score gap			0.009***				0.007***				-0.001	
			(0.002)				(0.002)				(0.002)	
Residual score gap				0.028***				0.024***				0.011***
				(0.002)				(0.002)				(0.002)
Observed score levels			-0.039***	-0.033***			-0.059***	-0.053***			-0.032***	-0.030***
			(0.002)	(0.002)			(0.002)	(0.002)			(0.002)	(0.002)
N	42,091	39,563	40,145	39,393	30,213	28,749	29,177	28,656	20,997	20,213	20,499	20,163
Mean of Dep. Var. Prob(separation)	0.151				0.149				0.081			

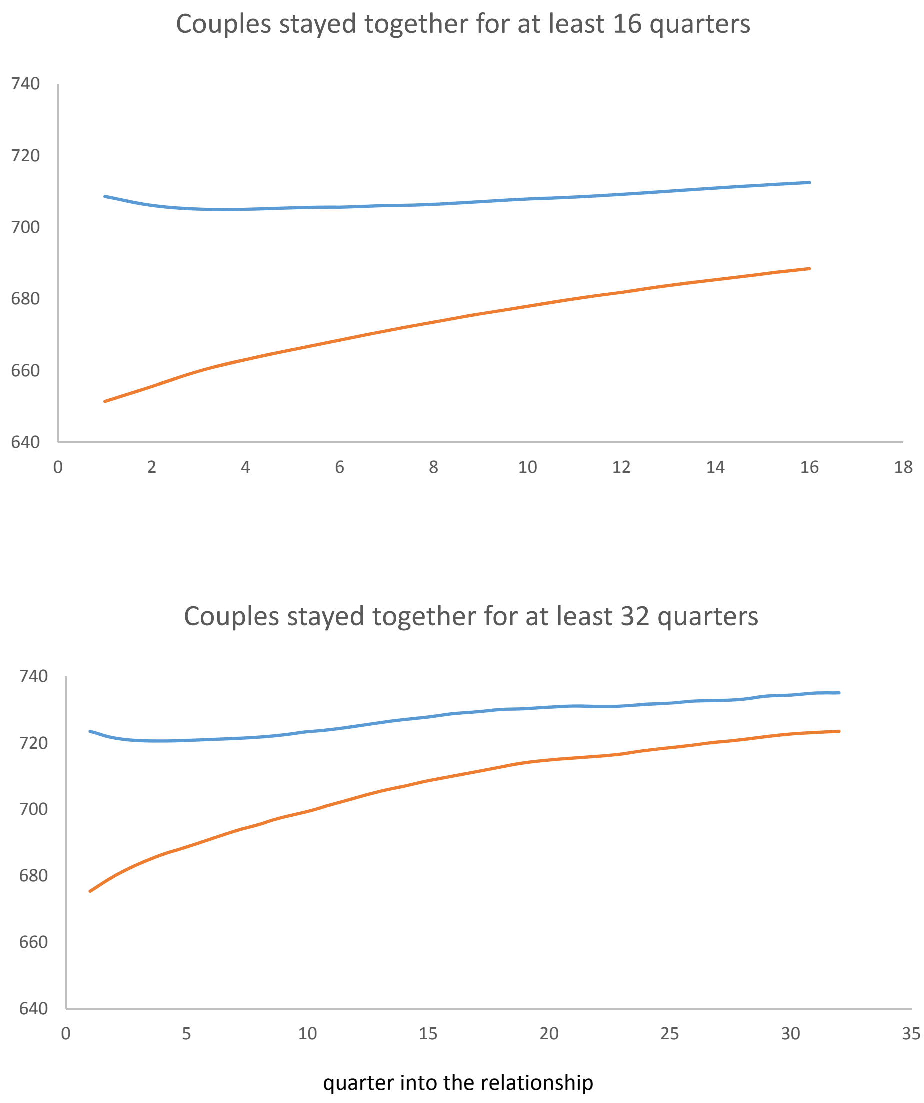
Note. The table reports marginal effect (dy/dx) estimates. Standard errors of estimated marginal effects are reported in parentheses. Each spouse's credit score is projected on eight lagged quarterly unemployment rate, personal bankruptcy rate, and delinquency rate of the county that individual lived before forming the relationship, a third-order age polynomial, and yearly dummies. Credit score are decomposed into a predicted component and a residual, using which the couple's predict and residual credit score averages and gaps are calculated. *, **, and *** denotes the estimate is statistically significant at the 90, 95, and 99 percent level, respectively. The estimations control for a third-order age polynomial of each spouse and the couple age differential; These estimations also control for differentials in socioeconomic characteristics, using differences of U.S. Census statistics of race, education, and income between the census block groups where each spouse lived prior to forming the relationship as proxies. These estimations also control for the share of divorced population where the couple lived as of the end of the first, second, and fourth year of their relationship, respectively. Finally, the estimations control for yearly and state fixed effects.

Figure 1: Credit Scores and Relationship Formation



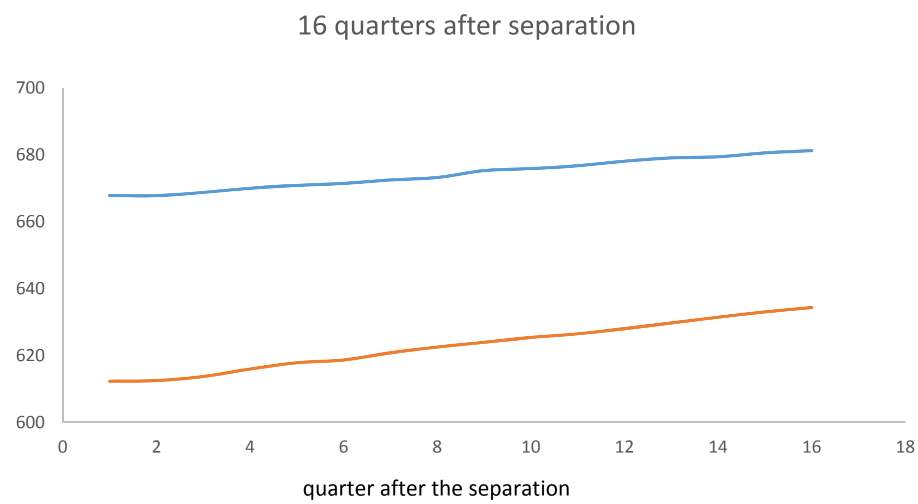
Source: Authors' estimation using the FRBNY Consumer Credit Panel / Equifax Data. Note: Single individuals of credit scores lower than 450 are the omitted group. The dotted lines plot the 95-percent confidence interval. The mean of the dependent variable is 12.5%.

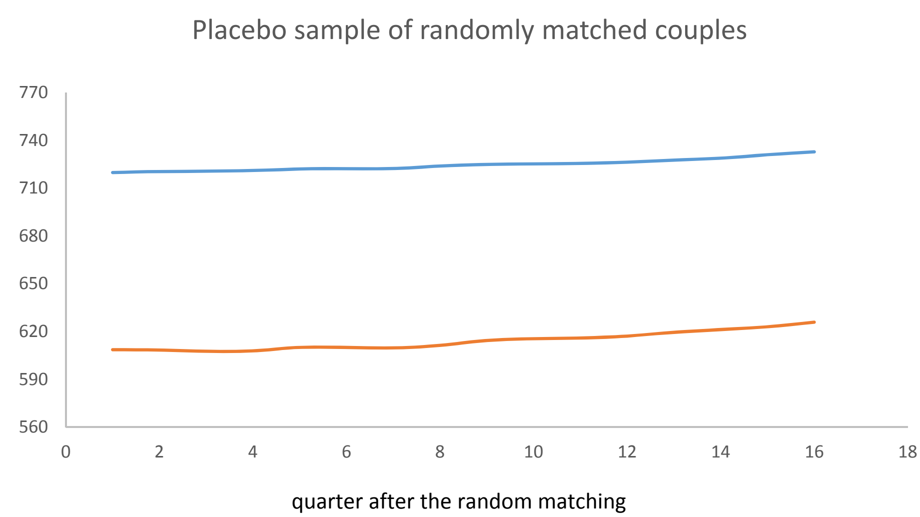
Figure 2: Dynamics of Credit Score Gaps of Lasting Relationships



Source: Authors' calculation using the FRBNY Consumer Credit Panel / Equifax Data.
Note: The blue curve shows the average credit score of the individuals who had the higher score of a couple as of the time of relationship formation. The orange curve shows the average credit score of the individuals who had the lower score of a couple as of the time of relationship formation. The series are estimated using the couples where both individuals are in the primary sample.

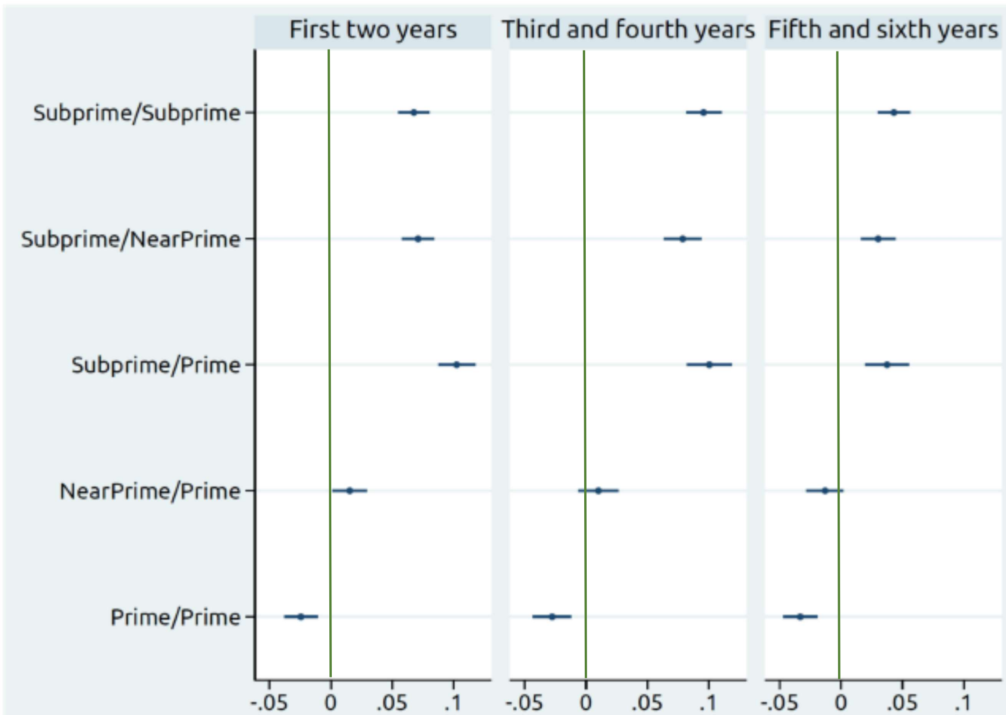
Figure 3: Dynamics of Credit Score Gaps of Separated and Placebo Relationships





Source: Authors' calculation using the FRBNY Consumer Credit Panel / Equifax Data.
Note: The blue curve shows the average credit score of the individuals who had the higher score of a couple as of the time of relationship formation. The orange curve shows the average credit score of the individuals who had the lower score of a couple as of the time of relationship formation. The series are estimated using the couples where both individuals are in the primary sample. The placebo sample is constructed by randomly matching individuals of the primary sample, respecting the age and age differential restrictions imposed on the identified couples.

Figure 4: Likelihood of Relationship Dissolution by Credit Levels and Gaps



Note: Estimated marginal effects of relationship dissolving in the first two, third and fourth, and fifth and sixth years. Relationships involving spouses both with Nearprime credit scores are the omitted group.

Appendix

A Robustness Analysis

Here, we provide a series of analyses to further support our conclusion that the initial match quality in the credit scores of those in committed relationships is predictive of subsequent separations. To begin with, recognizing the possibility that the pairs identified by our algorithm may not be true, committed relationships, we are particularly concerned that the pairs with especially large differences in their credit scores are not true couples but are driving the results. Noting that estimates from the Current Population Survey find that less than 5 percent of two-person households who satisfy our restrictions are not in committed relationships, we discuss the results of a trimming exercise where we drop 5 percent of the observations with the largest initial score differences.[23] The results of this exercise show that the estimates in Table 7 are essentially unchanged.

The remaining results from additional robustness checks are reported in Table A1 and discussed below.

Using Credit Score Percentiles

First, we consider whether using credit score values in the match quality estimate may be affecting the interpretation of our results. Credit scores are developed to reflect a rank ordering of individuals' credit risk and so the distance between the two partners' credit scores may be measured by differences in their percentile ranking rather than by differences in the values of the credit scores. As shown in columns 1 and 2 in the upper panel of Table A1, the level and match quality effects of credit scores remain pronounced when credit score percentiles, rather than levels, are used. For example, all else equal, a one standard deviation increase of initial percentile differences corresponds to a 3.1 and 1.2 percentage points higher likelihood of separation by the end of the fourth and sixth years of the relationship, respectively.[24] This similarity should not surprise us as we obtain similar patterns in the degree of assortative matching and measuring within-couple differences in credit scores when percentiles are used (see Table 6).

Inferring the Timing of Relationship Formation

Two methodological choices affect how we link pre-relationship credit score conditions to relationship outcomes: assigning the quarter in which a couple begins living together as the marker of the start of a committed relationship and the length of the period during which a couple cannot be observed to live together before they start their relationship. On the first choice, we assume in the baseline analysis that credit score differentials measured in the quarter in which a couple first begins living together appropriately measure the match quality in credit scores. It may be that partners' financial behaviors influence one another prior to when they begin living together. Moreover, it could also be that one partner in a couple does not immediately change his or her mailing address precisely at the same time that the committed

[23]These results are available upon request.

[24]The results of this and other robustness test specifications regarding relationship outcomes as of the end of the second year are also qualitatively unchanged relative to the baseline specification.

relationship begins. To address this concern, instead of the relationship formation quarter Q, we use each partner's credit score eight quarters before Q. These results, shown in columns 3 and 4 of Table A1, are qualitatively similar to the baseline estimates of separation hazards.

On the second choice, it may be that some couples use separate addresses even after forming a committed relationship, such as would be the case where an employee who is sent to work at a different location on a long-term basis than his or her home city. In this case, our baseline algorithm may mistakenly treat their moving back to the same location as a new relationship. To address this concern, we adjust the algorithm to limit the two partners sharing the same address to 16, rather than 8, quarters before Q, the quarter in which the relationship forms. These results, shown in columns 5 and 6, again are qualitatively similar to the baseline estimates.[25]

Including Couples Involving Nonprimary Sample Individuals

We now examine whether the baseline results are sensitive to adding couples that include non-primary sample individuals. Allowing such couples being included in the analysis substantially enlarged the sample of couples.[26] However, because the first time we observe the non-primary sample spouse is when that individual began living at the same address with a primary sample individual, we do not observe the census block group the non-primary sample individuals lived prior to forming the relationship. As a result, we cannot control for Z^{gap}, the approximated socioeconomic and demographic gaps, as we did in the baseline analysis. The results are reported in columns 7 and 8 in the lower panel of the table, and the implied marginal effect estimates are very similar to those in the baseline analysis.

Focusing on Couples with Joint Credit Accounts

One possible concern with our sample is that our algorithm of identifying committed relationships may actually include pairs of individuals who are not in a cohabiting or marital relationship. To rule out this possibility, we focus our attention on couples sharing joint debt accounts as non-committed couples are unlikely to jointly apply for credit cards, mortgages, or auto loans and share the financial obligations of repayment. We show, in columns 9 and 10, that for such couples the initial credit score match quality remains predictive for future separations.

B A Survey-Based Trustworthiness Measure and Credit Scores

B.1 Comparing These Two Measures of Trustworthiness

We turn to survey data to shed some light on the association between credit scores and personal character. The 2000 Social Capital Community Survey (SCCS) is a survey of 375 to 1,500 adults in 41 communities that ask respondents trust-related questions, such as "whether most

[25]The sample sizes are smaller in this specification because we do not observe all individuals of the baseline sample couples up to 16 quarters prior to relationship formation.

[26]However, because other variables are more likely to be left blank for non-primary sample individuals, the numbers of observations used in the estimations are smaller than those reported in Table 1.

people can be trusted or you can't be too careful."[27] While the questions targeted on trust attitude, Glaeser et al. (2000) and Sapienza et al. (2013) show that answers to these questions are more indicative of one's own trustworthiness rather than how much the individual gives to others. Moreover, people who interact with more trustworthy people and businesses in their neighborhoods tend to build a more trusting attitude (Alesina and Ferrara, 2002).

Accordingly, we construct an index of trustworthiness as the weighted percentage of survey respondents in a community who reply that "most people can be trusted."[28] Values of this index range from 36 to 66, and have a standard deviation of 8 across the 38 communities that are also covered in the CCP data. This index is highly correlated with the average credit scores of the corresponding areas in the CCP, with a Spearman rank correlation above 0.70, as shown in Figure B1. Moreover, this correlation persists after controlling for each survey community's level of income. As seen in column 1 of Table B1, when we project the average community credit scores onto the trustworthiness index, the estimated coefficient is highly significant, and the R-squared is above 50 percent. Regressing average credit score on the logarithm of the community's median income (not shown) yield an R-squared 0.26, only half of the R-squared in column 1. Finally, adding the log(median income) to the specification leaves the coefficient on the trustworthiness index little changed (column 2).

We then analyze whether previous exposure to trustworthiness is predictive of one's current credit score.[29] We construct a subsample of individuals in the CCP living in the communities surveyed by the SCCS for at least three years before moving out from these areas. We limit our analysis to those living in a community for at least three years to ensure sufficient time for any neighborhood effects to manifest. We then follow these individuals and estimate the correlation with contemporaneous credit scores. We find that previous exposure to neighborhood social trust statistically predicts contemporaneous credit scores. We acknowledge that this correlation is not a test for *causal* neighborhood effects as individuals likely sort into neighborhoods by their credit scores and other individual-level characteristics. Finally, we find suggestive evidence regarding how these survey-base indicator of trustworthiness help predict relationship outcomes and this analysis is discussed in the appendix.

B.2 The Survey-Based Measures and Relationship Outcomes

We explore the extent to which such a survey-based indicator of trustworthiness help predict relationship outcomes. First, we estimate a multinominal logistic model to test whether this indicator is associated with survey participants' marital status. We consider three marital status—never married, divorced or separated, and currently married, with the latter being the base outcome group. The regression also takes advantage of the race, gender, education, and income data included in the SCCS. The results, presented in Tables B2, demonstrate a

[27]These trust-related questions significantly overlap with those from the World Value Survey and the General Social Survey, the most frequently used data for trust and trustworthiness research.

[28]To be sure, whether answers to such survey questions measure trusting behavior or trustworthiness remains unsettled. For example, Fehr et al. (2003) challenges the results in Glaeser et al. (2000).

[29]See Guiso et al. (2004); Brown et al. (2008) for evidence on enduring neighborhood effects on financial market participation.

strong association between the trustworthiness indicator and marital status.[30] The odds ratios estimated from the entire sample (column 1) indicate that individuals who reply that "most people can be trusted" enjoyed a 10 percent lower odds of being never married and 15 percent lower odds of becoming divorced or separated. Estimating the model for three different age groups reassures a robust association over the lifecycle. Note that lower trustworthiness does not imply a higher never-married odds among younger survey participants (column 2), as many people of this age group are in the middle of the partner selection process.

We then identify a subsample of the identified couples in the CCP data that lived in areas covered by SCCS and merge in the SCCS trustworthiness index. We then estimate a model similar to Equation (3) to test whether exposure to certain levels of neighborhood trustworthiness may have a bearing on relationship outcomes.[31] We first focus on the CCP individuals who have always lived in these areas—the non-movers. As presented in column 1 in Table B3, higher local trustworthiness is negatively associated with the likelihood of relationship dissolving in the first six years. The estimate is statistically significant, with the standard error clustered at the SCCP survey area (SAMP) level. Interestingly, when adding couples' credit scores at relationship formation to the model, the trustworthiness index coefficient becomes smaller and statistically insignificant (column 2), consistent with our suggestive interpretation of credit scores containing information pertinent to one's trustworthiness. We then repeat the analysis on the sample that moved out of the SCCS-covered areas—the movers. The estimates are qualitatively similar but smaller in magnitude and not statistically significant, consistent with a lasting but diminishing effect of past trustworthiness exposures.

[30] For the multinominal logistic analysis, we show the odds ratios instead of marginal effects to facilitate interpretation.

[31] To increase the sample size, we do not require both spouses to be in the primary sample of the CCP. Accordingly, unlike in Equation (3), we are not able to include $\theta_g Z^{gap,i}$ as a control.

Table A1: Robustness Tests

	Credit score percentiles		Primary sample individuals only scores 8 Qs before		Living sep. 16 Qs before	
	year 3 or 4	year 5 or 6	year 3 or 4	year 5 or 6	year 3 or 4	year 5 or 6
	(1)	(2)	(3)	(4)	(5)	(6)
$\frac{Initial\ diff}{100}$	0.031***	0.012***	0.018***	0.110***	0.194***	0.100***
	(0.002)	(0.002)	(0.002)	(0.024)	(0.021)	(0.032)
$\frac{Initial\ score}{100}$	-0.066***	-0.038***	-0.043***	-0.356***	-0.413***	-0.427***
	(0.003)	(0.002)	(0.002)	(0.028)	(0.025)	(0.038)
Controlling for						
Age polynomials	yes	yes	yes	yes	yes	yes
Age diff.	yes	yes	yes	yes	yes	yes
Current community char.	yes	yes	yes	yes	yes	yes
Community div. pop. share	yes	yes	yes	yes	yes	yes
Orig. community char. diff.	yes	yes	yes	yes	yes	yes
Yearly FE	yes	yes	yes	yes	yes	yes
State FE	yes	yes	yes	yes	yes	yes
N	29,188	20,518	28,345	19,974	14,516	8,618

	Including non-primary sample consumers: Whole sample		Including non-primary sample consumers: Having joint accounts ever	
	year 3 or 4	year 5 or 6	year 3 or 4	year 5 or 6
	(7)	(8)	(9)	(10)
$\frac{Initial\ diff}{100}$	0.248***	0.161***	0.061***	0.026***
	(0.002)	(0.003)	(0.004)	(0.005)
$\frac{Initial\ score}{100}$	-0.431***	-0.414***	-0.375***	-0.324***
	(0.003)	(0.004)	(0.005)	(0.005)
Controlling for				
Age polynomials	yes	yes	yes	yes
Age diff.	yes	yes	yes	yes
Current community char.	yes	yes	yes	yes
Community div. pop. share	yes	yes	yes	yes
Orig. community char. diff.	No	No	No	No
Yearly FE	yes	yes	yes	yes
State FE	yes	yes	yes	yes
N	1,195,160	820,103	755,210	588,395

Note. The table reports marginal effects (dy/dx) of the logistic regressions associated with a one-standard-deviation change of credit score differentials and average levels at the time of relationship formation. Standard errors are reported in parentheses. *** denotes the estimate is statistically significant at the 99 percent level. The estimations control for a third-order age polynomial of each spouse and the couple age differential; The estimations control for U.S. Census statistics of race, education, income, and share of divorced population of the census block group where the couple lived as of the end of the first, second, and fourth year of their relationship, respectively, for the three sets of results. For results of column 1–8, the estimations also control variables the differentials of these statistics between the census block groups where each spouse lived prior to forming the relationship. Finally, the estimations control for yearly and state fixed effects. See text for details of each of the robustness test specifications.

Table B1: Survey Based Trustworthiness Index and Credit Scores

	Contemporary correlations		Long-term influences
	Community average credit score		Individual credit score
	(1)	(2)	(3)
Trustworthiness Index	1.57***	1.42***	
	(0.25)	(0.21)	
Trustworthiness Index (community lived 3 years ago)			0.61***
			(0.03)
Log(median income)		0.36***	0.42***
		(0.09)	(0.01)
R-squared	0.51	0.65	0.008
N	38	38	340,303

Note. The trustworthiness index is defined as 100 times the share of respondents who replied that "most people can be trusted" to the 2000 Social Capital Community Survey (SCCS) for each community surveyed, using the SCCS weights. Median income is the 2000 U.S. Census statistics of the corresponding community. Results reported in column 3 concern the individuals who lived in the communities covered by the survey for at least three years before moving out of these areas. The individuals' credit scores were taken three years after the move.

Table B2: Survey-Based Trust and Trustworthiness Indicator and Marital Status: A Multinominal Logistic Analysis

	Odds ratios relative to "currently married = 1" associated with the trusting/trustworthy indicator			
	(1)	(2)	(3)	(4)
	All	35 and younger	36–50	51–65
Never married = 1	0.900***	0.962	0.873*	0.746**
	(0.036)	(0.049)	(0.061)	(0.093)
Divorced or separated = 1	0.853***	0.797**	0.883**	0.831**
	(0.036)	(0.072)	(0.054)	(0.067)
Controlling for				
Race	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes
Gender	Yes	Yes	Yes	Yes
Income	Yes	Yes	Yes	Yes
Age	Yes	No	No	No
City fixed effects	Yes	Yes	Yes	Yes
N	21,791	8,777	8,472	4,542

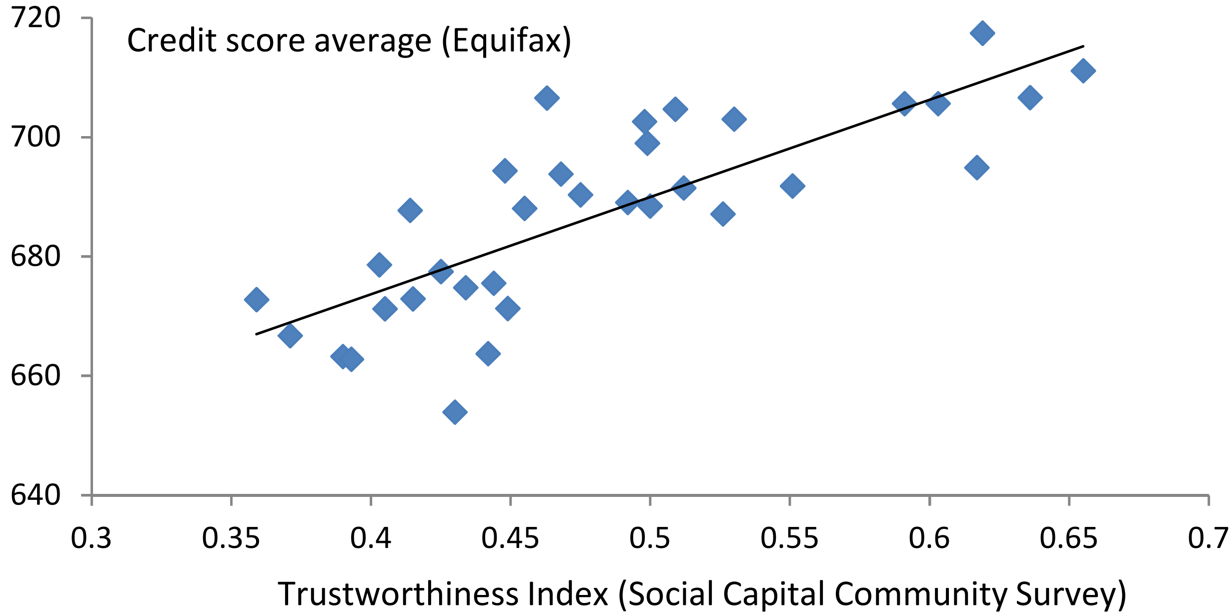
Note. Report odds ratios of a measure of trust and trustworthiness on marital status estimated with a multinominal logistic regression. The trusting-trustworthy indicator is set to equal to one for the respondents who replied that "most people can be trusted" to the 2000 Social Capital Community Survey (SCCS). The likelihood of being never married or divorced (separated) is estimated relative to that of currently being married. The model includes controls of race, educational attainments, gender, income, age, and city fixed effects. * denotes the odds ratio estimate is statistically significant at the 90 percent level, and *** denotes the estimate is statistically significant at the 99 percent level. The differential between the two coefficients of the bottom row is also statistically significant at the 99 percent level.

Table B3: Neighborhood Trust-Trustworthiness Exposure and Relationship Dissolution within Six Years

	Non-movers		Movers	
	(1)	(2)	(3)	(4)
Trustworthiness index	-0.176**	-0.097	-0.132	-0.045
	(0.082)	(0.098)	(0.094)	(0.117)
Credit score		-0.110***		-0.115***
(normalized by sample std.)		(0.003)		(0.003)
Controls				
Age polynomials	Yes	Yes	Yes	Yes
Age differential	Yes	Yes	Yes	Yes
Primary-sample individuals education and income proxies	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
N	112,429	112,429	44,107	44,107
Mean of Dep. Var. Prob(separation)	0.495		0.461	

Note. Report marginal effects (dy/dx) estimated from a logit model of CCP sample relationship dissolution on a local trustworthiness index and other controls. The index is calculated for each survey area as the share of respondents who replied that "most people can be trusted" to the 2000 Social Capital Community Survey (SCCS). The CCP sample includes the couples living in the areas covered by the SCCS. Non-movers refer to the CCP individuals who stayed in these areas during the entire sample periods, whereas movers refer to those who moved out of the SCCS areas. ** denotes the estimate is statistically significant at the 95 percent level, and *** denotes the estimate is statistically significant at the 99 percent level.

Figure B1: Credit Scores and Trustworthiness



Source: Authors' calculation using the FRBNY Consumer Credit Panel / Equifax Data and the Social Capital Community Survey.