Assessing racial and educational segmentation in large marriage markets

Edoardo Ciscato

KU Leuven
FEB - Department of Economics

August 29, 2023

Motivation

Positive Assortative Mating (PAM):

- Observed in many dimensions: age, race, education, religion... (Schwartz, 2013)
- Linked with economic inequality (Fernández and Rogerson, 2001; Chiappori et al., 2017) and intergenerational transmission of preferences (Bisin and Verdier, 2000; Dohmen et al., 2012)

Two main concurrent explanations

1. Household complementarities (Becker, 1973)
   $\rightarrow$ PAM associated with welfare gains from optimal sorting
2. Search frictions (Oppenheimer, 1988; Mortensen, 1988)
   $\rightarrow$ PAM results from constraints on meeting opportunities

Motivation

Positive Assortative Mating (PAM):

- Observed in many dimensions: age, race, education, religion... (Schwartz, 2013)
- Linked with economic inequality (Fernández and Rogerson, 2001; Chiappori et al., 2017) and intergenerational transmission of preferences (Bisin and Verdier, 2000; Dohmen et al., 2012)

Two main concurrent explanations

1. Household complementarities (Becker, 1973)
   $\rightarrow$ PAM associated with welfare gains from optimal sorting
2. Search frictions (Oppenheimer, 1988; Mortensen, 1988)
   $\rightarrow$ PAM results from constraints on meeting opportunities

Research question and contributions

To what extent are the observed patterns of PAM shaped by search frictions vs complementarities?

This paper provides three contributions (see literature)

1. Introduces a dynamic model of marriage market steady-state equilibrium with heterogeneous frictions, sorting, and endogenous separations and rematches
2. Discusses the joint identification of the match surplus and frictions with relationship duration data
3. Estimates the model with 2009-2013 US data on romantic relationships and shows that
   - Racial homogamy can almost entirely be explained by frictions
   - Educational homogamy would be half as strong without frictions

Research question and contributions

To what extent are the observed patterns of PAM shaped by search frictions vs complementarities?

This paper provides three contributions (see literature)

1. Introduces a dynamic model of marriage market steady-state equilibrium with heterogeneous frictions, sorting, and endogenous separations and rematches
2. Discusses the joint identification of the match surplus and frictions with relationship duration data
3. Estimates the model with 2009-2013 US data on romantic relationships and shows that
   - Racial homogamy can almost entirely be explained by frictions
   - Educational homogamy would be half as strong without frictions

Research question and contributions

To what extent are the observed patterns of PAM shaped by search frictions vs complementarities?

This paper provides three contributions (see literature)

1. Introduces a dynamic model of marriage market steady-state equilibrium with heterogeneous frictions, sorting, and endogenous separations and rematches
2. Discusses the joint identification of the match surplus and frictions with relationship duration data
3. Estimates the model with 2009-2013 US data on romantic relationships and shows that
   - Racial homogamy can almost entirely be explained by frictions
   - Educational homogamy would be half as strong without frictions

Research question and contributions

To what extent are the observed patterns of PAM shaped by search frictions vs complementarities?

This paper provides three contributions (see literature)

1. Introduces a dynamic model of marriage market steady-state equilibrium with heterogeneous frictions, sorting, and endogenous separations and rematches
2. Discusses the joint identification of the match surplus and frictions with relationship duration data
3. Estimates the model with 2009-2013 US data on romantic relationships and shows that
    - Racial homogamy can almost entirely be explained by frictions
    - Educational homogamy would be half as strong without frictions

Model

Model outline: environment

Discrete time, large population, stationary environment, overlapping generations (see model equations)

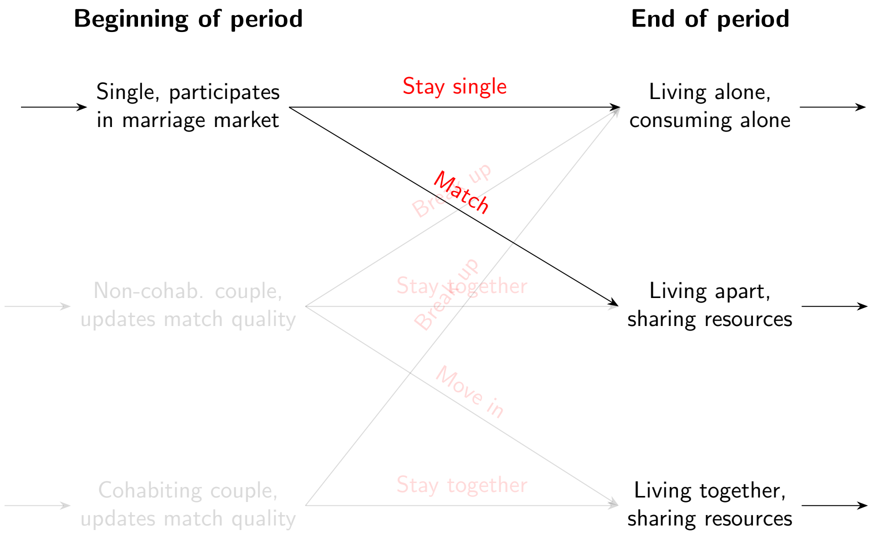
- **Individuals** are heterogeneous
  - Inborn exogenous traits: gender, birthyear, race, and education
  - Aging: marriage market entry at age 19, death at 84
  - Relationship status: agents can be single or in a couple at
- **Couples** are characterized by
  - Partners' types i and j
  - Living arrangement a: non-cohabiting ($a = NC$) and cohabiting relationships ($a = C$)
  - Time-invariant match quality k: low ($k = L$) or high ($k = H$)
  - A bargaining parameter $\theta \in (0, 1)$ set at the start of the relationship

Model outline: environment

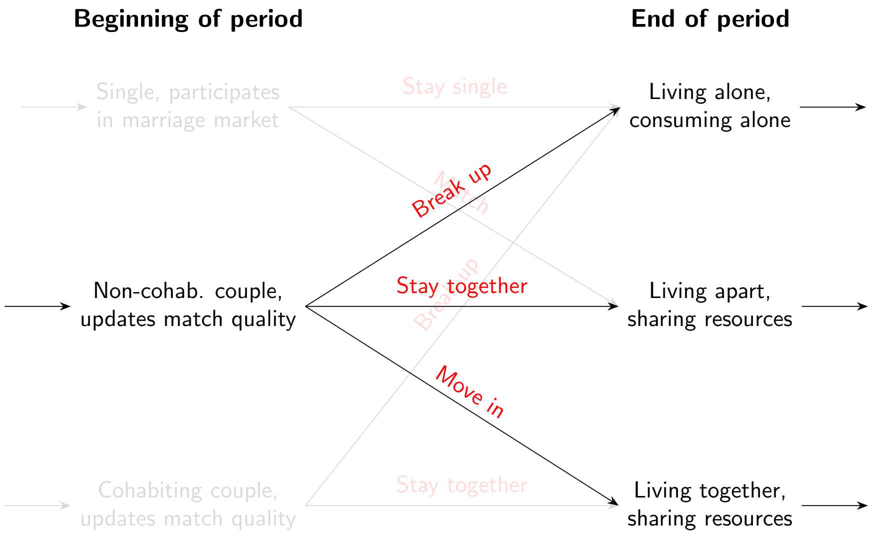
Discrete time, large population, stationary environment, overlapping generations (see model equations)

- **Individuals** are heterogeneous
  - Inborn exogenous traits: gender, birthyear, race, and education
  - Aging: marriage market entry at age 19, death at 84
  - Relationship status: agents can be single or in a couple at
- **Couples** are characterized by
  - Partners' types i and j
  - Living arrangement a: non-cohabiting ($a = NC$) and cohabiting relationships ($a = C$)
  - Time-invariant match quality k: low ($k = L$) or high ($k = H$)
  - A bargaining parameter $\theta \in (0, 1)$ set at the start of the relationship

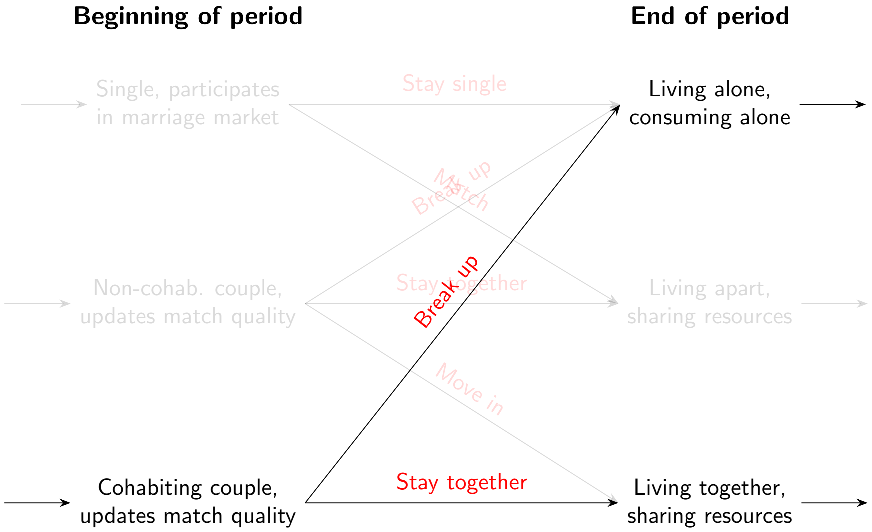
Model outline: matching and breaking up



Model outline: matching and breaking up



Model outline: matching and breaking up



Relationship gains

Total match surplus:

$$S_{i,j,a,k} + z = H_{i,j,a,k} + z + \mathcal{C}_{i,j,a,k} - V^m_{i,0} - V^f_{0,j} \tag{1}$$

- $H_{i,j,a,k}$: per-period relationship gains, *model primitive*, depends on input complementarities
- z: match-quality component updated in every period, distributed logistically
- $\mathcal{C}_{i,j,a,k}$: couple's continuation value, depends on exp. match surplus and prospects of continuing the relationship (and moving in together)
- $V^m_{i,0}$ and $V^f_{0,j}$: partners' reservation values, depend on their marriage market prospects

1. Couples break up if $S_{i,j,a,k} + z$ turns negative
2. Total gains $S_{i,j,a,k} + z$ are shared through Nash bargaining

Relationship gains

Total match surplus:

$$S_{i,j,a,k} + z = H_{i,j,a,k} + z + \mathcal{C}_{i,j,a,k} - V^m_{i,0} - V^f_{0,j} \qquad (1)$$

- $H_{i,j,a,k}$: per-period relationship gains, *model primitive*, depends on input complementarities
- z: match-quality component updated in every period, distributed logistically
- $\mathcal{C}_{i,j,a,k}$: couple's continuation value, depends on exp. match surplus and prospects of continuing the relationship (and moving in together)
- $V^m_{i,0}$ and $V^f_{0,j}$: partners' reservation values, depend on their marriage market prospects

1. Couples break up if $S_{i,j,a,k} + z$ turns negative
2. Total gains $S_{i,j,a,k} + z$ are shared through Nash bargaining

Marriage markets

Competitive bipartite one-to-one matching game: single women solve

$$\max\left\{\max_i\left\{\theta_{i,j}\left(\bar{S}_{i,j}-\Lambda_{i,j}\right)+\varepsilon_i^f\right\},\ \ 0\right\} \tag{2}$$

- $\Lambda_{i,j,a,k}$: meeting costs, *model primitive*
- $\bar{S}_{i,j}=\mathbb{E}_k[S_{i,j,NC,k}|i,j]$: exp. match surplus, persistent match quality k not known yet
- $\theta_{i,j}$: woman j's surplus share when matched with man of type i
- ε^f: woman j's random component of meeting costs

The sharing rule θ adjusts s.t.

1. All agents solve problem (2)
2. The number of matched agents of given type does not exceed the current (endogenous) supply

Marriage markets

Competitive bipartite one-to-one matching game: single women solve

$$\max\left\{\max_i\left\{\theta_{i,j}\left(\bar{S}_{i,j}-\Lambda_{i,j}\right)+\varepsilon_i^f\right\},\ \ 0\right\} \tag{2}$$

- $\Lambda_{i,j,a,k}$: meeting costs, *model primitive*
- $\bar{S}_{i,j}=\mathbb{E}_k[S_{i,j,NC,k}|i,j]$: exp. match surplus, persistent match quality k not known yet
- $\theta_{i,j}$: woman j's surplus share when matched with man of type i
- ε^f: woman j's random component of meeting costs

The sharing rule θ adjusts s.t.

1. All agents solve problem (2)
2. The number of matched agents of given type does not exceed the current (endogenous) supply

Econometrics

Matching function

We estimate the *matching function* implied by our setup

$$\frac{MF_{i,j}}{\sqrt{MF_{i,\emptyset}MF_{\emptyset,j}}} = \exp\left(\frac{1}{2}(\bar{S}_{i,j} - \Lambda_{i,j})\right) \tag{3}$$

provided we observe the matching outcome on the lhs

Setup: i, j and $a \in \{NC, C\}$ are observed for all singles and couples

Problem: how to separately identify $\bar{S}_{i,j}$ and $\Lambda_{i,j}$ as they both depend on the same observables (i.e. partners' race, education, age...)?

Matching function

We estimate the *matching function* implied by our setup

$$\frac{MF_{i,j}}{\sqrt{MF_{i,\emptyset}MF_{\emptyset,j}}} = \exp\left(\frac{1}{2}(\bar{S}_{i,j} - \Lambda_{i,j})\right) \tag{3}$$

provided we observe the matching outcome on the lhs

Setup: i, j and $a \in \{NC, C\}$ are observed for all singles and couples

Problem: how to separately identify $\bar{S}_{i,j}$ and $\Lambda_{i,j}$ as they both depend on the same observables (i.e. partners' race, education, age...)?

Identification without time-invariant match quality

When relationships do not differ by time-invariant quality k

$$\frac{MF_{i,j}}{\sqrt{MF_{i,\emptyset} MF_{\emptyset,j}}} = \exp\left(\frac{1}{2}(S_{i,j,NC} - \Lambda_{i,j})\right) \tag{4}$$

With data on separations, we infer $S_{i,j,NC}$ from

$$\Pr\{\text{breaking up}|i,j,NC\} = \Pr\{S_{i,j,NC} + z < 0|i,j,NC\} \tag{5}$$

Key condition: we must observe incumbent couples before their stakes change, i.e. before partners move in together, get married, have children...

Identification with time-invariant match quality

In the presence of time-invariant match quality k, "soulmates" are more likely to endure temporary negative shocks

Identification of the match surplus $S_{i,j,NC,k}$ and the initial probability of drawing k, $\pi_{i,j,k}$, is obtained from

$$\left\{\Pr\{\text{breaking up}|i, j, NC, \text{stayed together for } d \text{ years}\}\right\}_{d=1}^{D} \quad (6)$$

with a panel of length $D \geq 2K - 1$

Data

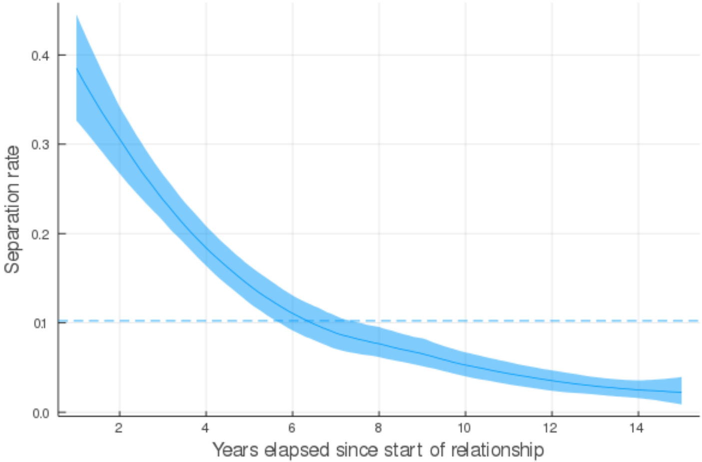
Data

How Couples Meet and Stay Together (2009-2013)

- Covers all romantic relationships (also non-cohabiting)
- A match is defined as two individuals who start dating
- Retrospective information on relationship duration
- Separations observed in annual follow-up surveys
- Race: White, Black, Hispanic and Other
- Education: high school diploma or less, some college, college degree or more

See descriptive stats

Figure: Separation rate by relationship duration



Notes. Duration is measured as years elapsed since a couple started dating (without necessarily cohabiting).

Results

Parameterization

- The meeting cost function is parameterized as

$$\Lambda_{i,j} = x_{i,j}^{\intercal} \lambda \tag{7}$$

  where $x_{i,j}$ is a vector of basis functions
- People older than 63 no longer search for a mate
- Per-period match gains are parameterized as

$$H_{i,j,a,k} = x_{i,j}^{\intercal} \delta + \zeta_a + \eta_k \tag{8}$$

  where ζ_a and η_k are fixed effects for living arrangements a and match quality k

Table: GMM estimates

	Meeting cost $\Lambda_{i,j}$ (1)	Match gains $H_{i,j,a,k}$ (2)
Same age group	-2.79	0.26
	(0.42)	(0.10)
Same race	-3.90	0.02
	(0.63)	(0.06)
Same education	-0.43	0.10
	(0.55)	(0.05)

Notes. Standard errors in parentheses.

Other results

- See all parameter estimates
  - Frictions increasing in age, relationship gains decreasing in age
  - Higher relationship gains for college graduates
  - Stronger frictions for non-white women
  - Strong selection on unobservables over the first five years
- See validation exercise
  - Low costs, high gains: partners coming from same or met through friends
  - Low costs, high gains: partners coming from same town or high school; met through family
  - High costs, high gains: online meetings, parties
  - High costs, low gains: workplace
- Robustness checks
- Model fit

Counterfactual analysis

Decomposing the matching function

Holding the supplies of available singles $\tilde{n}^m$ and $\tilde{n}^f$ fixed, we calculate

1. The random matching outcome
2. The matching outcome if singles sorted only on *meeting costs*
3. The matching outcome if singles sorted only on *match gains*

See full comparative statics between steady-state equilibria

Table: Homogamy patterns among new couples - data vs counterfactuals

		Counterfactuals		
	Data	Random assignment (1)	No match gains (2)	No meeting costs (3)
% of same-race couples	77.5	39.1	73.1	39.5
% of same-educ couples	36.8	24.0	32.0	28.3

Decomposing the matching function

Holding the supplies of available singles $\tilde{n}^m$ and $\tilde{n}^f$ fixed, we calculate

1. The random matching outcome
2. The matching outcome if singles sorted only on *meeting costs*
3. The matching outcome if singles sorted only on *match gains*

See full comparative statics between steady-state equilibria

Table: Homogamy patterns among new couples - data vs counterfactuals

		Counterfactuals		
	Data	Random assignment (1)	No match gains (2)	No meeting costs (3)
% of same-race couples	77.5	39.1	73.1	39.5
% of same-educ couples	36.8	24.0	32.0	28.3

Conclusion

Main takeaways

- Structural approach to address the question *who meets whom*
  - It uses information from both relationship formation and dissolution contained in panel data
  - It draws from both search-and-matching and the frictionless Choo-Siow literature
  - It allows us to study meetings in large nationally representative datasets
- Application: measuring educational and racial segmentation
  - Strong evidence of racial segmentation, weak evidence of racial complementarities
  - Strong evidence of educational complementarities, relatively less important role played by educational segmentation
  - Removing racial segmentation increases overall instability but can be beneficial to minorities with a narrow market

Appendix

Literature ◀

- Econometrics of matching models: Choo and Siow (2006); Dupuy and Galichon (2014); Bruze et al. (2015); Choo (2015); Galichon et al. (2019); Galichon and Salanié (2022)...
- Search and matching applied to marriage markets: Shimer and Smith (2000); Wong (2003); Greenwood et al. (2016); Goussé et al. (2017); Shephard (2018)...
- Measurement of PAM with nationally representative samples: Fryer Jr (2007); Chiappori et al. (2017); Eika et al. (2019); Chiappori et al. (2020)...
- Measurement of matching preferences vs meeting opportunities: Fisman et al. (2006, 2008); Nielsen and Svarer (2009); Hitsch et al. (2010); Belot and Francesconi (2013); Kirkebøen et al. (2021)...

Literature ◀

- Econometrics of matching models: Choo and Siow (2006); Dupuy and Galichon (2014); Bruze et al. (2015); Choo (2015); Galichon et al. (2019); Galichon and Salanié (2022)...
- Search and matching applied to marriage markets: Shimer and Smith (2000); Wong (2003); Greenwood et al. (2016); Goussé et al. (2017); Shephard (2018)...
- Measurement of PAM with nationally representative samples: Fryer Jr (2007); Chiappori et al. (2017); Eika et al. (2019); Chiappori et al. (2020)...
- Measurement of matching preferences vs meeting opportunities: Fisman et al. (2006, 2008); Nielsen and Svarer (2009); Hitsch et al. (2010); Belot and Francesconi (2013); Kirkebøen et al. (2021)...

Cohabiting couples ◀

- At the beginning of the period, partners draw *match quality* z from a logistic distribution
- A match generates per-period gains $H_{i,j,C,k} + z$
- Gains are shared through Nash bargaining

$$\max_v \Bigg\{ \theta \ln\Big(v + \mathcal{C}^m_{i,j,C,k}(\theta) - V^m_{i,0} \Big) + \\ + (1-\theta)\ln\Big(H_{i,j,C,k} + z + \mathcal{C}^f_{i,j,a',k}(\theta) - V^f_{0,j} \Big) \Bigg\} \tag{9}$$

  - $\mathcal{C}^m_{i,j,C,k}(\theta)$ and $\mathcal{C}^f_{i,j,C,k}(\theta)$ denote the *continuation values*
  - $V^m_{i,0}$ and $V^f_{0,j}$ denote the *reservation values*

Cohabiting couples ◀

The male partner gets a fraction θ of the *total surplus* $S_{i,j,C,k}$

$$\underbrace{v^m_{i,j,C,k}(\theta, z) + \mathcal{C}^m_{i,j,C,k} - V^m_{i,0}}_{\text{Male partner's share}} = \\ = \theta \Bigg[\underbrace{H_{i,j,C,k} + z + \mathcal{C}_{i,j,C,k} - V^m_{i,0} - V^f_{0,j}}_{\text{Match surplus}} \Bigg] \tag{10}$$

The probability that a cohabiting couple breaks up is

$$1 - \alpha_{i,j,C,k} = \mathbb{E}_z\big[\mathbb{1}\{S_{i,j,C,k} + z < 0\}\big] = \big[1 + \exp(S_{i,j,C,k})\big]^{-1} \tag{11}$$

Cohabiting couples ◀

The male partner gets a fraction θ of the *total surplus* $S_{i,j,C,k}$

$$\underbrace{v^m_{i,j,C,k}(\theta, z) + \mathcal{C}^m_{i,j,C,k} - V^m_{i,0}}_{\text{Male partner's share}} = \\ = \theta \left[\underbrace{H_{i,j,C,k} + z + \mathcal{C}_{i,j,C,k} - V^m_{i,0} - V^f_{0,j}}_{\text{Match surplus}} \right] \quad (10)$$

The probability that a cohabiting couple breaks up is

$$1 - \alpha_{i,j,C,k} = \mathbb{E}_z \left[\mathbb{1}\{S_{i,j,C,k} + z < 0\} \right] = \left[1 + \exp(S_{i,j,C,k}) \right]^{-1} \quad (11)$$

Non-cohabiting couples ◀

- At the beginning of the period, partners draw match-quality shocks (z_{NC}, z_C) from a logistic distribution
- Moving in together has a cost κ
- Couples choose

$$\max\{\underbrace{S_{i,j,NC,k} + z_{NC}}_{\text{Don't move in}},\ \underbrace{S_{i,j,C,k} - \kappa + z_C}_{\text{Move in}},\ \underbrace{0}_{\text{Break up}}\} \tag{12}$$

The probability that a cohabiting couple breaks up is

$$1 - \alpha^{NC}_{i,j,NC,k} - \alpha^{C}_{i,j,NC,k} = \left[1 + \exp(S_{i,j,NC,k}) + \exp(S_{i,j,C,k} - \kappa)\right]^{-1} \tag{13}$$

NB: can be extended to allow for correlation between z_{NC} and z_C

Non-cohabiting couples ◀

- At the beginning of the period, partners draw match-quality shocks (z_{NC}, z_C) from a logistic distribution
- Moving in together has a cost κ
- Couples choose

$$\max\{\underbrace{S_{i,j,NC,k} + z_{NC}}_{\text{Don't move in}},\ \underbrace{S_{i,j,C,k} - \kappa + z_C}_{\text{Move in}},\ \underbrace{0}_{\text{Break up}}\} \tag{12}$$

The probability that a cohabiting couple breaks up is

$$1 - \alpha^{NC}_{i,j,NC,k} - \alpha^{C}_{i,j,NC,k} = \left[1 + \exp(S_{i,j,NC,k}) + \exp(S_{i,j,C,k} - \kappa)\right]^{-1} \tag{13}$$

NB: can be extended to allow for correlation between z_{NC} and z_C

Singles and mate search ◀

- A single woman draws logistic taste shocks $\{\varepsilon_i^f\}_{i\in\mathcal{I}}$
- She chooses between a partner of type i and staying single

$$\bar{V}_j^f = \max\left\{\max_i\left\{(1-\theta_{i,j})\left(\bar{S}_{i,j}-\Lambda_{i,j}\right)+\varepsilon_i^f\right\},\ 0\right\} \qquad (14)$$

  - $\bar{S}_{i,j}$ is the expected match surplus before learning $k\in\{L,H\}$

$$\bar{S}_{i,j} = (1-\pi_{i,j})S_{i,j,NC,L} + \pi_{i,j}S_{i,j,NC,H} \qquad (15)$$

  - $\Lambda_{i,j}$ is an exogenous cost couples must pay to engage in a relationship
  - $\theta_{i,j}$ is the share of surplus demanded by a male partner of type i
- If she does not find a partner, she can look again in the next period

$$V_{0,j}^f = \beta\mathbb{E}_{j'}V_{0,j'}^f + \beta\mathbb{E}_{j'}\bar{V}_{j'}^f \qquad (16)$$

Singles and mate search ◀

- A single woman draws logistic taste shocks $\{\varepsilon_i^f\}_{i \in \mathcal{I}}$
- She chooses between a partner of type i and staying single

$$\bar{V}_j^f = \max\left\{\max_i \left\{(1-\theta_{i,j})\left(\bar{S}_{i,j} - \Lambda_{i,j}\right) + \varepsilon_i^f\right\},\ 0\right\} \tag{14}$$

  - $\bar{S}_{i,j}$ is the expected match surplus before learning $k \in \{L, H\}$

$$\bar{S}_{i,j} = (1-\pi_{i,j})S_{i,j,NC,L} + \pi_{i,j}S_{i,j,NC,H} \tag{15}$$

  - $\Lambda_{i,j}$ is an exogenous cost couples must pay to engage in a relationship
  - $\theta_{i,j}$ is the share of surplus demanded by a male partner of type i
- If she does not find a partner, she can look again in the next period

$$V_{0,j}^f = \beta \mathbb{E}_{j'} V_{0,j'}^f + \beta \mathbb{E}_{j'} \bar{V}_{j'}^f \tag{16}$$

Singles and mate search ◀

- A single woman draws logistic taste shocks $\{\varepsilon_i^f\}_{i\in\mathcal{I}}$
- She chooses between a partner of type i and staying single

$$\bar{V}_j^f = \max\left\{\max_i \left\{(1-\theta_{i,j})\left(\bar{S}_{i,j} - \Lambda_{i,j}\right) + \varepsilon_i^f\right\},\ \ 0\right\} \tag{14}$$

  - $\bar{S}_{i,j}$ is the expected match surplus before learning $k \in \{L, H\}$

$$\bar{S}_{i,j} = (1-\pi_{i,j})S_{i,j,NC,L} + \pi_{i,j}S_{i,j,NC,H} \tag{15}$$

  - $\Lambda_{i,j}$ is an exogenous cost couples must pay to engage in a relationship
  - $\theta_{i,j}$ is the share of surplus demanded by a male partner of type i
- If she does not find a partner, she can look again in the next period

$$V_{0,j}^f = \beta\mathbb{E}_{j'}V_{0,j'}^f + \beta\mathbb{E}_{j'}\bar{V}_{j'}^f \tag{16}$$

Marriage markets ◀

- Demand for partners is given by problem (14)
- Supply is fixed: $\tilde{n}_i^m$ ($\tilde{n}_j^f$) is the number of single men (women) of type i (j) at the beginning of each period
- $\theta_{i,j}$ adjusts so that demand and supply meet

The number of new matches between types (i, j) is given by

$$MF_{i,j} = \exp\left(\frac{1}{2}(\bar{S}_{i,j} - \Lambda_{i,j})\right)\sqrt{MF_{i,\emptyset}MF_{\emptyset,j}} \tag{17}$$

$$MF_{\emptyset,j} = \tilde{n}_i^f - \sum_i MF_{i,j} \geq 0 \tag{18}$$

$$MF_{i,\emptyset} = \tilde{n}_i^m - \sum_j MF_{i,j} \geq 0 \tag{19}$$

Marriage markets ◀

- Demand for partners is given by problem (14)
- Supply is fixed: $\tilde{n}_i^m$ ($\tilde{n}_j^f$) is the number of single men (women) of type i (j) at the beginning of each period
- $\theta_{i,j}$ adjusts so that demand and supply meet

The number of new matches between types (i, j) is given by

$$MF_{i,j} = \exp\left(\frac{1}{2}(\bar{S}_{i,j} - \Lambda_{i,j})\right)\sqrt{MF_{i,\emptyset}MF_{\emptyset,j}} \tag{17}$$

$$MF_{\emptyset,j} = \tilde{n}_i^f - \sum_i MF_{i,j} \geq 0 \tag{18}$$

$$MF_{i,\emptyset} = \tilde{n}_i^m - \sum_j MF_{i,j} \geq 0 \tag{19}$$

Steady-state search equilibrium ◀

Key exogenous parameters:

- Household production function H
- Meeting cost function Λ
- Initial distribution of time-invariant match quality π
- Population supplies (e.g. gender ratio, % of college graduates)

Key endogenous (and stationary) quantities:

- Reservation values and match surplus
- Number of new and incumbent matches, separations, and singles
- Nash-bargaining weight function θ and sharing rule

Steady-state search equilibrium ◀

Key exogenous parameters:

- Household production function H
- Meeting cost function Λ
- Initial distribution of time-invariant match quality π
- Population supplies (e.g. gender ratio, % of college graduates)

Key endogenous (and stationary) quantities:

- Reservation values and match surplus
- Number of new and incumbent matches, separations, and singles
- Nash-bargaining weight function θ and sharing rule

Table: Descriptive statistics - estimation sample ◀

	All		Singles		Partnered	
	Women	Men	Women	Men	Women	Men
A. Sample composition						
# of respondents	1,465	1,408	301	244	1,164	
# of observations	3,833	3,776	301	244	3,532	
In a relationship (%)	65.2	71.4				
B. Age and relationship characteristics						
Age (years)	36.9	37.1	42.5	39.9	33.9	36.0
Cohabiting (%)					76.8	
Duration (years)					6.8	

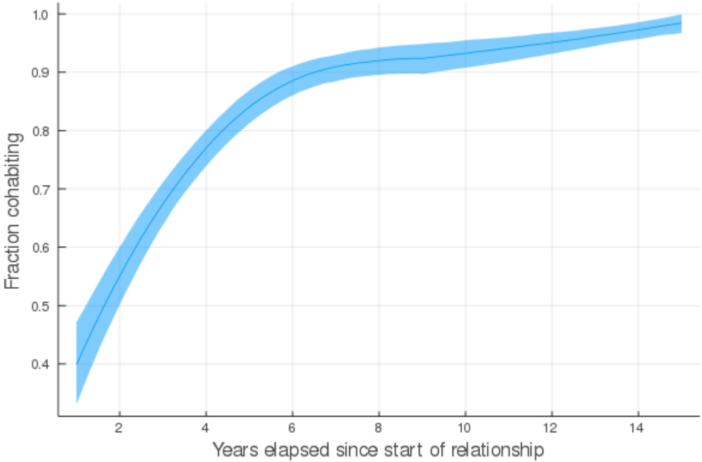
Notes. By construction of the sample, age is capped at 62 and relationship duration at 15.

Table: Descriptive statistics - estimation sample ◀

	All		Singles		Partnered	
	Women	Men	Women	Men	Women	Men
C. Race (%)						
White	68.7	70.7	62.5	73.2	72.1	69.7
Black	13.2	10.4	19.9	9.9	9.7	10.6
Hispanic	11.3	13.4	12.0	12.4	11.0	13.9
Other	6.7	5.5	5.7	4.5	7.3	5.8
D. Education (%)						
High-school dropout	8.3	6.6	9.4	3.4	7.8	7.9
High-school diploma	28.3	28.0	34.7	28.7	24.8	27.7
Some college	31.0	35.1	29.1	40.9	32.0	32.8
Bachelor's degree	22.8	20.9	18.4	19.2	25.1	21.7
Graduate degree	9.6	9.3	8.4	7.9	10.3	9.9

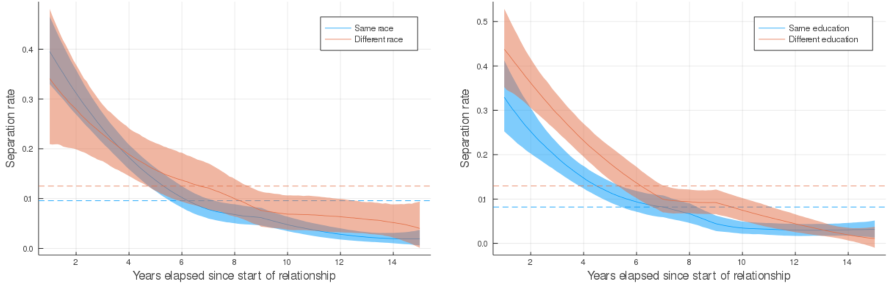
Notes. In panel C, each column details the racial composition of the sample in 2009 conditional on gender and relationship status, with every column summing to 100. Similarly, panel D details the educational composition of the sample in 2009, again with every column summing to 100.

Figure: Share cohabiting by relationship duration ◀



Notes. Duration is measured as years elapsed since a couple started dating (without necessarily cohabiting).

Figure: Separation rate by relationship duration and type of relationship ◀



Notes. Duration is measured as years elapsed since a couple started dating (without necessarily cohabiting).

Table: Match and separation rates - education ◀

	Match rate (%)		Separation rate (%)	
	Women	Men	Women	Men
Education				
High-school dropout	25.5	36.5	21.8	9.1
High-school diploma	9.1	22.7	12.1	15.0
Some college	23.9	24.1	17.4	17.3
Bachelor's degree	21.4	24.2	8.8	7.6
Graduate degree	24.5	19.7	6.4	8.6

Notes. Match rates are calculated as the ratio between the number of couples formed in a year and the number of singles at the beginning of the year. Separation rates correspond to the fraction of couples that break up in a given year.

Table: Match and separation rates - race ◀

	Match rate (%)		Separation rate (%)	
	Women	Men	Women	Men
Race				
White	21.2	22.4	11.7	12.0
Black	13.2	30.3	13.0	18.1
Hispanic	16.4	15.0	20.0	13.0
Other	19.4	45.6	13.9	12.9

Notes. Match rates are calculated as the ratio between the number of couples formed in a year and the number of singles at the beginning of the year. Separation rates correspond to the fraction of couples that break up in a given year.

Table: Match and separation rates - age ◀

	Match rate (%)		Separation rate (%)	
	Women	Men	Women	Men
Age				
19-25	32.4	31.7	24.1	24.0
26-34	37.2	35.1	8.5	11.4
35-42	23.6	33.3	12.1	7.1
43-50	8.5	14.4	9.1	15.1
51-62	6.7	9.8	13.4	12.6

Notes. Match rates are calculated as the ratio between the number of couples formed in a year and the number of singles at the beginning of the year. Separation rates correspond to the fraction of couples that break up in a given year.

Table: Match characteristics conditional on meeting circumstances ◀

	Same race (%)	Same education (%)	Man's age (years)	Woman's age (years)
Overall	82.5	55.2	29.1	27.1
	-	-	-	-
Same town	87.2	61.2	25.6	23.8
	(1.6)	(3.2)	(0.0)	(0.0)
Same high school	86.5	62.0	23.6	22.2
	(9.5)	(4.1)	(0.0)	(0.0)
Same college	84.5	83.3	22.4	20.9
	(44.8)	(0.0)	(0.0)	(0.0)
Met online	78.8	52.9	33.2	30.8
	(34.7)	(49.8)	(0.0)	(0.0)
Met through friends	80.5	59.2	27.8	26.0
	(28.5)	(6.6)	(0.2)	(0.3)
Met through family	83.6	47.5	28.9	27.5
	(54.8)	(1.6)	(82.6)	(65.0)
Met through neighbors	84.1	64.1	27.1	24.9
	(54.2)	(6.3)	(3.6)	(2.1)

Notes. Every line provides the % of racially and educationally homogamous couples, along with the male and female partner's age, conditional on specific meeting circumstances. In parentheses, we report p-values (multiplied by 100) corresponding to the probabilities that a given estimate is equal to the corresponding estimate for the entire estimation sample.

Table: Match characteristics conditional on meeting circumstances ◀

	Same race (%)	Same education (%)	Man's age (years)	Woman's age (years)
Overall	82.5	55.2	29.1	27.1
	-	-	-	-
Met through coworkers	78.2	45.3	30.5	28.4
	(14.6)	(0.1)	(2.0)	(3.5)
Met at work	80.6	48.5	31.6	28.3
	(44.6)	(2.9)	(0.0)	(4.9)
Met at church	87.8	62.1	24.0	22.0
	(19.6)	(25.7)	(0.0)	(0.0)
Met on vacation	90.4	46.7	27.5	27.4
	(23.2)	(43.0)	(45.5)	(89.3)
Met in a bar	88.2	53.2	32.6	31.4
	(9.0)	(66.7)	(0.0)	(0.0)
Met in an association	81.6	61.9	26.3	24.6
	(88.8)	(42.6)	(5.7)	(5.9)
Met at a party	76.6	63.1	29.2	27.3
	(8.1)	(5.1)	(92.3)	(78.0)

Notes. Every line provides the fraction of racially and educationally homogamous couples, along with the male and female partner's age, conditional on specific meeting circumstances. In parentheses, we report p-values (multiplied by 100) corresponding to the probabilities that a given estimate is equal to the corresponding estimate for the entire estimation sample.

Table: GMM estimates - age ◀

	Meeting cost $\Lambda_{i,j}$ (1)	Match gains $H_{i,j,a,k}$ (2)
Constant	14.38	0.35
	(1.00)	(0.14)
Woman's age	1.07	-0.10
	(0.16)	(0.03)
Man's age	0.66	-0.08
	(0.18)	(0.02)
Same age group	-2.79	0.26
	(0.42)	(0.10)

Notes. Standard errors in parentheses.

Table: GMM estimates - race ◀

	Meeting cost $\Lambda_{i,j}$ (1)	Home production $\bar{H}_{i,j}$ (2)
Black (woman)	1.58	-0.06
	(0.85)	(0.08)
Hispanic (woman)	0.16	-0.24
	(0.93)	(0.08)
Other (woman)	1.51	-0.18
	(1.24)	(0.11)
Black (man)	-1.60	0.00
	(0.87)	(0.08)
Hispanic (man)	1.45	-0.05
	(1.20)	(0.07)
Other (man)	-2.27	0.15
	(1.28)	(0.16)
Same race	-3.90	0.02
	(0.63)	(0.06)

Notes. Standard errors in parentheses. White is the reference category for racial groups.

Table: GMM estimates - education ◀

	Meeting cost $\Lambda_{i,j}$ (1)	Match gains $H_{i,j,a,k}$ (2)
High-school dropout (woman)	-2.63	0.22
	(1.11)	(0.11)
Some college (woman)	-2.01	0.09
	(0.75)	(0.06)
Bachelor's degree (woman)	-1.28	0.22
	(0.93)	(0.09)
Graduate degree (woman)	-0.73	0.39
	(1.08)	(0.15)
High-school dropout (man)	1.51	0.32
	(1.08)	(0.18)
Some college (man)	-0.26	-0.00
	(0.64)	(0.05)
Bachelor's degree (man)	0.97	0.19
	(0.77)	(0.08)
Graduate degree (man)	1.95	0.10
	(1.16)	(0.08)
Same education	-0.43	0.10
	(0.55)	(0.05)

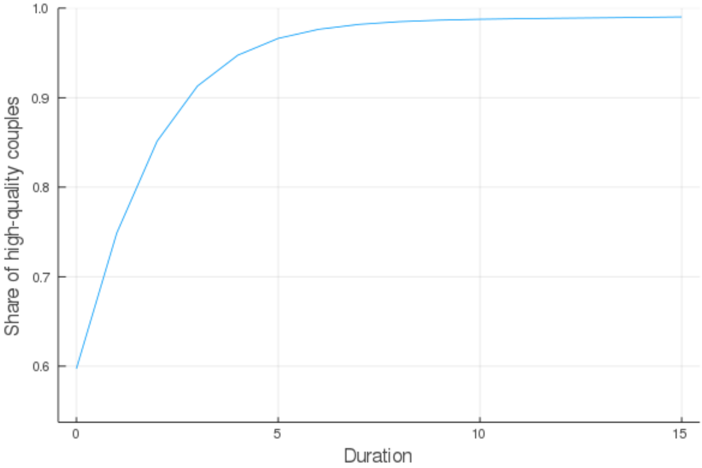
Notes. Standard errors in parentheses. High school diploma is the reference category for education.

Table: GMM estimates - additional parameters ◄

	Estimate
Gains from cohabiting (ζ_C)	0.37
	(0.06)
Gains from high quality (η_H)	0.69
	(0.18)
Cost of moving in (κ)	2.57
	(0.20)
Prob. of high-quality match (π)	0.61
	(0.07)

Notes. Standard errors in parentheses.

Figure: Share of high-quality relationships by duration ◀



Notes. A high-quality relationship is characterized by a high time-invariant match quality component k.

Interpreting the structural estimates ◀

- For a couple l with traits (i, j) in our dataset, $y_{i,j,l}$ is a variable describing meeting circumstances $\rightarrow$ we run the regression

$$y_{i,j,l} = \beta_0 + \beta_1 \Lambda_{i,j} + \beta_2 \bar{S}_{i,j} + \xi_l, \tag{20}$$

- For a couple l with traits (i, j) in living arrangment a and duration d in our dataset, $y_{i,j,a,d,l}$ is a variable describing match quality $\rightarrow$ we run the regression

$$y_{i,j,a,d,l} = \beta_4 + \beta_5 S_{i,j,a,\bar{k}} + \zeta_l, \tag{21}$$

Interpreting the structural estimates ◀

- For a couple l with traits (i, j) in our dataset, $y_{i,j,l}$ is a variable describing meeting circumstances → we run the regression

$$y_{i,j,l} = \beta_0 + \beta_1 \Lambda_{i,j} + \beta_2 \bar{S}_{i,j} + \xi_l, \tag{20}$$

- For a couple l with traits (i, j) in living arrangment a and duration d in our dataset, $y_{i,j,a,d,l}$ is a variable describing match quality → we run the regression

$$y_{i,j,a,d,l} = \beta_4 + \beta_5 S_{i,j,a,\bar{k}} + \zeta_l, \tag{21}$$

Table: Correlation between meeting costs and observed meeting circumstances ◀

	OLS estimates	
	Meeting costs $\Lambda_{i,j}$	Expected match surplus $\bar{S}_{i,j}$
Same town	-3.20	-2.40
	(0.40)	(1.50)
Same high school	-2.80	-1.59
	(0.35)	(1.28)
Same college	-1.14	3.94
	(0.26)	(0.98)
Met online	0.99	2.63
	(0.34)	(1.24)
Met through friends	-0.69	2.36
	(0.50)	(1.84)
Met through family	-0.33	-2.12
	(0.38)	(1.40)
Met through neighbors	-0.09	0.20
	(0.28)	(1.05)

Notes. Standard errors in parentheses. The variables in the first column are all dummies, and all coefficients are multiplied by 100. Both regressors, $\Lambda_{i,j}$ and $\bar{S}_{i,j}$, are scaled by the variance of the bliss shock $Var[z]$.

Table: Correlation between meeting costs and observed meeting circumstances ◄

	OLS estimates	
	Meeting costs $\Lambda_{i,j}$	Expected match surplus $\bar{S}_{i,j}$
Met through coworkers	0.90	-0.04
	(0.40)	(1.47)
Met at work	0.99	-1.38
	(0.40)	(1.50)
Met at church	-0.56	0.07
	(0.21)	(0.80)
Met on vacation	0.27	0.10
	(0.15)	(0.55)
Met in a bar	0.36	-0.32
	(0.29)	(1.07)
Met in an association	0.02	0.21
	(0.17)	(0.64)
Met at a party	0.62	4.18
	(0.32)	(1.18)

Notes. Standard errors in parentheses. The variables in the first column are all dummies, and all coefficients are multiplied by 100. Both regressors, $\Lambda_{i,j}$ and $\bar{S}_{i,j}$, are scaled by the variance of the bliss shock $Var[z]$.

Table: Correlation between match surplus and observed match characteristics ◀

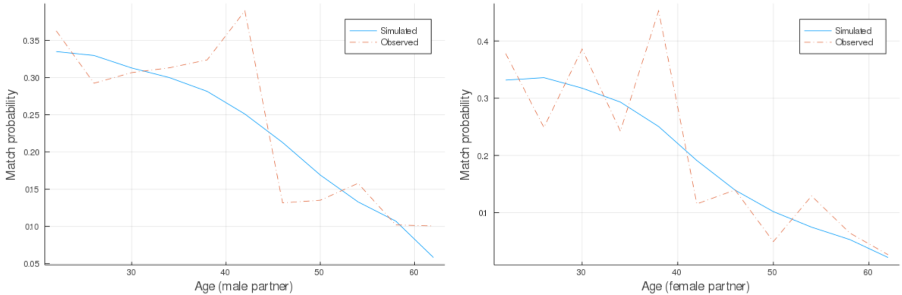
	Weighted OLS estimates
	Match surplus $S_{i,j,a,k}$
Legally married	10.83
	(0.68)
Relationship approved by parents	5.02
	(0.61)
Male partner earns more	1.66
	(0.74)
Female partner earns more	-0.31
	(0.65)
Unemployed (respondent)	-1.04
	(0.33)
Relationship satisfaction	0.06
	(0.01)
Log household income	0.10
	(0.01)

Notes. Standard errors in parentheses. $S_{i,j,a,k}$ is scaled by the variance of the bliss shock $Var[z]$. All variables in the first column are dummies, and all coefficients are multiplied by 100, with the exception of relationship satisfaction and log-household income. Relationship satisfaction is measured on a five-level Likert scale ranging from "poor" to "excellent".

Robustness checks ◀

- Three educational and two racial groups
- Differences in gains from cohabitation for homogamous vs mixed couples
- Differences in match-quality volatility for homogamous vs mixed couples
- Differences in odds of high-quality matches for homogamous vs mixed couples
- Nested logit version

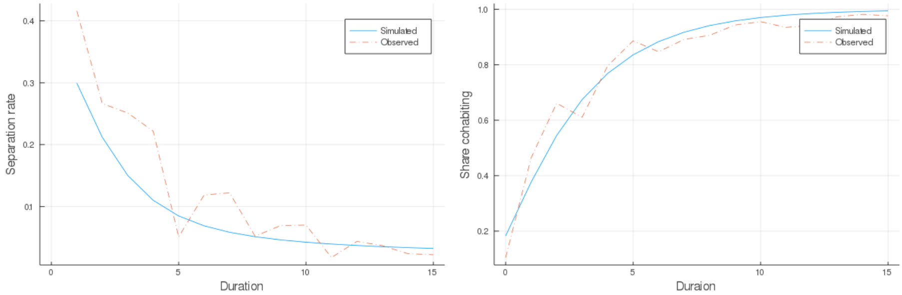
Figure: Age and duration profiles - model vs data ◀



(a) (b)

Notes. Match rates correspond to the odds of finding a partner over the next year. Solid lines represent simulated moments, dashed lines represent empirical moments.

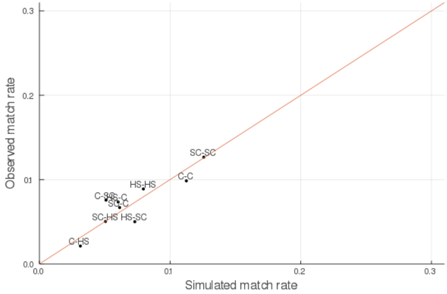
Figure: Age and duration profiles - model vs data ◀



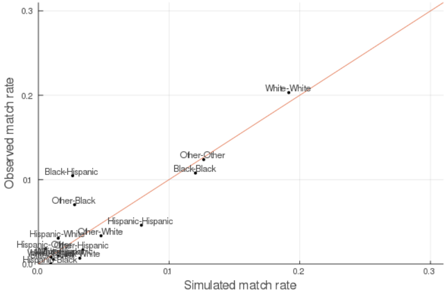
(a) (b)

Notes. Separation rates corresponds to the odds that a couple breaks up over the next year. Solid lines represent simulated moments, dashed lines represent empirical moments.

Figure: Match rates - model vs data ◀



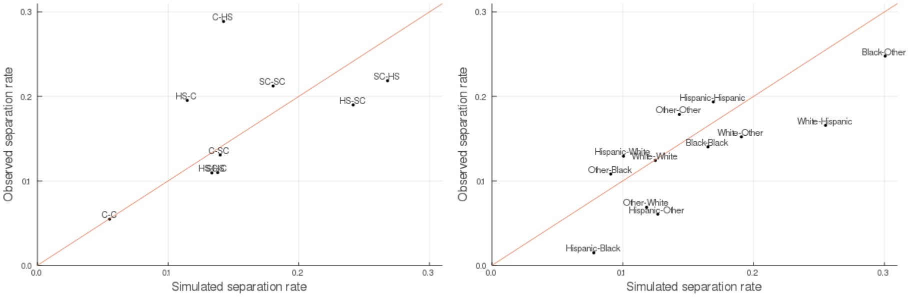
(a)



(b)

Notes. Simulated match rates conditional on partners' education are on the x-axis, the corresponding empirical moments are on the y-axis. Match rates are given by the number of yearly matches between a man of a given type and a woman of given type, divided by the square root of the product of the number of male and female singles of the corresponding types. Points that lie on the 45-degree line indicate a perfect fit.

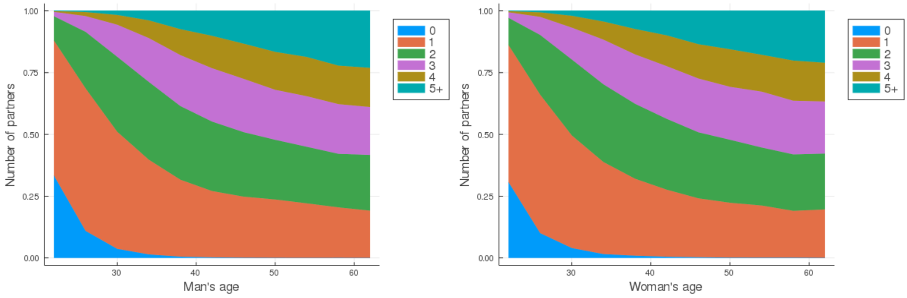
Figure: Separation rates - model vs data ◀



(a) (b)

Notes. Simulated separation rates conditional on partners' education are on the x-axis, the corresponding empirical moments are on the y-axis.

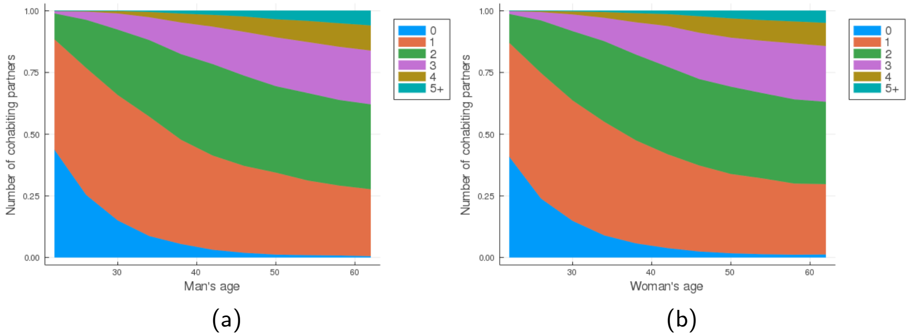
Figure: Number of partners by age and gender ◀



(a) (b)

Notes. The different panels show the simulated distribution of the total number of partners by age, gender, and type of relationship.

Figure: Number of partners by age and gender ◀



Notes. The different panels show the simulated distribution of the total number of cohabiting partners by age, gender, and type of relationship.

Comparative statics between steady-state equilbria ◀

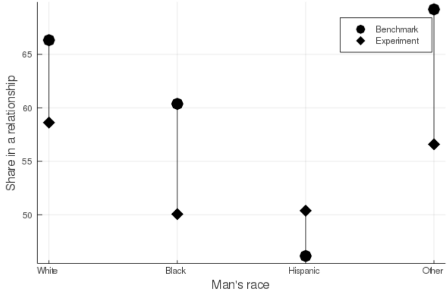
Counterfactual steady-states

- We hold the household production function constant
    - I. We remove additional meeting costs for different-race pairs
    - II. We remove additional meeting costs for different-education pairs
- We hold the structure of meeting costs constant
    - III. We remove match complementarities for same-race pairs
    - IV. We remove match complementarities for different-education pairs

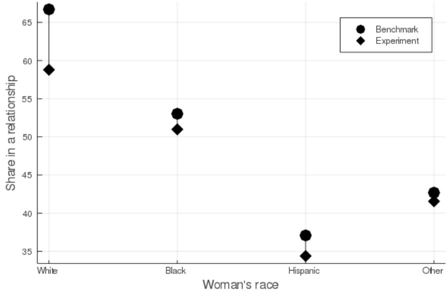
Table: Counterfactual experiments ◄

		No segmentation		No match compl	
	Benchmark	Exp. I Race	Exp. II Education	Exp. III Race	Exp. IV Education
% in a relationship (men)	63.1	56.5	61.3	61.9	60.4
% in a relationship (women)	60.2	54.0	58.5	59.0	57.4
Match rate (%, men)	20.7	27.7	21.8	20.3	19.8
Match rate (%, women)	17.8	24.2	18.8	17.5	17.0
Separation rate (%)	11.9	22.3	13.7	12.3	12.9
Share same-age couples (%)	23.4	25.5	23.8	23.5	23.6
Share same-race couples (%)	82.8	59.8	82.6	81.4	82.7
Share same-educ couples (%)	39.9	39.2	36.0	39.9	27.0
Share cohabiting (%)	73.0	60.4	70.2	72.4	71.6
Share high k (%)	90.5	85.7	89.5	90.2	90.0
Average z ($\times 100$)	14.9	32.4	17.6	15.5	16.2

Figure: Experiment I ◄

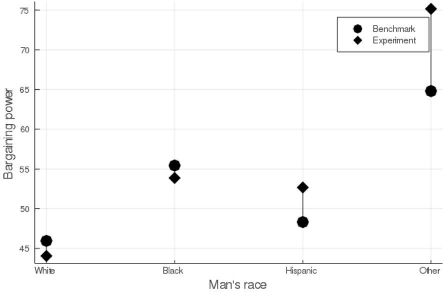


(a) Changes in % of partnered men

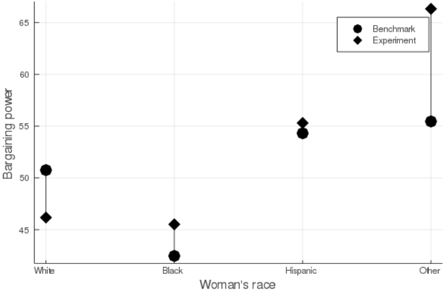


(b) Changes in % of partnered women

Figure: Experiment I ◀



(a) Changes in men's bargaining power $(1 - \theta)$



(b) Changes in women's bargaining power (θ)

Literature

Becker, G. S. (1973). A theory of marriage: Part i. Journal of Political Economy, 81(4):813–846.

Belot, M. and Francesconi, M. (2013). Dating preferences and meeting opportunities in mate choice decisions. Journal of Human Resources, 48(2):474–508.

Bisin, A. and Verdier, T. (2000). "beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits. The Quarterly Journal of Economics, 115(3):955–988.

Bruze, G., Svarer, M., and Weiss, Y. (2015). The dynamics of marriage and divorce. Journal of Labor Economics, 33(1):123–170.

Chiappori, P.-A., Dias, M. C., Crossman, S., and Meghir, C. (2020). Changes in assortative matching and inequality in income: Evidence for the uk. Working paper 26933, National Bureau of Economic Research.

Chiappori, P.-A., Salanié, B., and Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium. American Economic Review, 107(8):2109–67.

Choo, E. (2015). Dynamic marriage matching: An empirical framework. Econometrica, 83(4):1373–1423.

Choo, E. and Siow, A. (2006). Who marries whom and why. Journal of Political Economy, 114(1):175–201.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2012). The intergenerational transmission of risk and trust attitudes. The Review of Economic Studies, 79(2):645–677.

Dupuy, A. and Galichon, A. (2014). Personality traits and the marriage market. Journal of Political Economy, 122(6):1271–1319.

Eika, L., Mogstad, M., and Zafar, B. (2019). Educational assortative mating and household income inequality. Journal of Political Economy, 127(6):2795–2835.

Fernández, R. and Rogerson, R. (2001). Sorting and long-run inequality. Quarterly Journal of Economics, 116(4):1305–1341.

Fisman, R., Iyengar, S. S., Kamenica, E., and Simonson, I. (2006). Gender differences in mate selection: Evidence from a speed dating experiment. The Quarterly Journal of Economics, 121(2):673–697.

Fisman, R., Iyengar, S. S., Kamenica, E., and Simonson, I. (2008). Racial preferences in dating. The Review of Economic Studies, 75(1):117–132.

Fryer Jr, R. G. (2007). Guess who's been coming to dinner? trends in interracial marriage over the 20th century. Journal of Economic Perspectives, 21(2):71–90.

Galichon, A., Kominers, S. D., and Weber, S. (2019). Costly concessions: An empirical framework for matching with imperfectly transferable utility. Journal of Political Economy, 127(6):2875–2925.

Galichon, A. and Salanié, B. (2022). Cupid's invisible hand: Social surplus and identification in matching models.

Goussé, M., Jacquemet, N., and Robin, J.-M. (2017). Marriage, labor supply, and home production. Econometrica, 85(6):1873–1919.

Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2016). Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation. American Economic Journal: Macroeconomics, 8(1):1–41.

Hitsch, G. J., Hortaçsu, A., and Ariely, D. (2010). Matching and sorting in online dating. American Economic Review, 100(1):130–63.

Kirkebøen, L., Leuven, E., and Mogstad, M. (2021). College as a marriage market. Working paper 28688, National Bureau of Economic Research.

Mortensen, D. T. (1988). Matching: Finding a partner for life or otherwise. American Journal of Sociology, 94:S215–S240.

Nielsen, H. S. and Svarer, M. (2009). Educational homogamy how much is opportunities? Journal of Human Resources, 44(4):1066–1086.

Oppenheimer, V. K. (1988). A theory of marriage timing. American Journal of Sociology, 94(3):563–591.

Schwartz, C. R. (2013). Trends and variation in assortative mating: Causes and consequences. Annual Review of Sociology, 39:451–470.

Shephard, A. (2018). Marriage market dynamics, gender, and the age gap. Working paper.

Shimer, R. and Smith, L. (2000). Assortative matching and search. Econometrica, 68(2):343–369.

Wong, L. Y. (2003). Structural estimation of marriage models. Journal of Labor Economics, 21(3):699–727.