Dynamic Marriage Matching: An Empirical Framework

Eugene Choo
University of Calgary

Matching Problems Conference,

June 5th 2012

Introduction

- Interested in rationalizing the marriage distribution of ‘who marries whom’ by age.
- To allow for dynamics in marriage and marital decisions.
- Empirically quantify the marital gains across gender and age.
  How important are dynamic considerations in marital decisions?
- Propose a dynamic version of the Becker-Shapley-Shubik model.
- Model rationalizes a new marriage matching function,
  $\boldsymbol{\mu} = \mathcal{G}(\boldsymbol{m}, \boldsymbol{f}; \boldsymbol{\Pi})$ where $\boldsymbol{\mu}$ is the distribution of new marriages,
  $\boldsymbol{m}$ and $\boldsymbol{f}$ are the vectors of available single men and women,
  and $\boldsymbol{\Pi}$ is a matrix of parameters, .

Contributions - *Dynamic Marriage Matching Function*

$$\mu_{ij} = \Pi_{ij}\sqrt{m_i f_j}\prod_{k=0}^{z_{ij}}\left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k}f_{j+k}}\right)^{\frac{1}{2}(\beta S)^k}. \quad (1)$$

- (i, j) denote the ages (or types) of males and females respectively.
- μ_{ij} is the number of observed new (i, j) matches,
  μ_{i0} is the number of i men who remained single and
  μ_{0j} is the number of j women who remained single
- m_i and f_j are the number of single type i men and j women respectively.
- discount factor is β, divorce rate is δ, survival probability $S = 1 - \delta$
- $z_{ij} = Z - \max(i, j)$, measures the maximum length of a marriage.

Contributions - *Dynamic Marriage Matching Function* continues

- Π_{ij} is the present discounted value of an (i, j) match relative to remaining single for the duration of the match.

$$\Pi_{ij} = \sum_{k=0}^{z_{ij}} (\beta S)^k \left[(\alpha_{ijk} + \gamma_{ijk}) - (\alpha_{i+k,0} + \gamma_{0,j+k})\right] - 2\kappa \qquad (2)$$

- α_{ijk} be the $k'th$ period marital output accrued to a type i male when married to a type j female today,
- similarly γ_{ijk} be the $k'th$ period marital output accrued to a type j female when married to a type i male,
- α_{i0} and γ_{0j} are the per-period utilities from remaining single for i type males and j type females respectively.
- κ is the geometric sum of Euler's constants.

Empirical Application

- Use model to analyze the fall in marital gains by age and gender between 1970 and 1990 in the US
- Show that dynamic component marital gains is large especially among the young.
- Ignoring dynamics severely unstate the drop in marital gains between 1970 and 1990 especially among young couples.

Literature

- Builds on frictionless Becker-Sharpley-Shubik transferable utility model of marriage
- Extends ideas in Choo and Siow (2006) and Choo and Siow (2005)
- Adopt the dynamic discrete choice framework of Rust (1987)
- Growing body of empirical work on marriage matching:
  Chiappori, Salanie and Weiss (2011),
  Chiappori, McCann and Nesheim (2009),
  Galichon and Salanie (2010),
  Echenique, Lee, Shum and Yenmez (2011),
  Ariely, Hortacsu and Hitsch (2006, 2010),
  Fox (2010).

The Model - Assumptions

- **State Variables:** Single individuals has two state variables:

1. (i, j) denote male and female's age when single, terminal age is Z.
2. $\boldsymbol{\epsilon}_{ig}$, is a $(Z+1)$ vector of i.i.d idiosyncratic payoffs specific to type i male individual, g ($\boldsymbol{\epsilon}_{jG}$ for type j female, G), unobserved to econometrician. Agents observe $\boldsymbol{\epsilon}$ at beginning of period.

- **Stationarity:** Single males and females, m_i and f_j $\forall i, j$ at each period taken as given.
- **Actions:** $a_{ig} \in \{0, 1, \ldots, Z\}$ (or a_{jG}) denote the action of a single type i male g (or single type j female G).

  If g (or G) chooses to remain single, $a_{ig} = 0$ (or $a_{jG} = 0$),
  else if g (or G) chooses to match with a type k spouse, $a_{ig} = k$ (or $a_{jG} = k$).

The Model - Assumptions continues

- **Exogenous Parameters:** discount factor is β, divorce rate δ, the survival probability $S = 1 - \delta$.
- Adopt **Dynamic Discrete Choice framework** of Rust(1987), maintain Rust's Additive Separability (AS) and Conditional Independence (CI).
- **Additive Separability (AS)** in utilities
  Utility function of a single male g decomposes to

$$v(a_{ig}, i, \boldsymbol{\epsilon}_{ig}) = v_a(i) + \epsilon_{iag},$$

  similarly utility function of a single female G takes the form,

$$w(a_{jG}, j, \boldsymbol{\epsilon}_{jG}) = w_a(j) + \epsilon_{jaG}.$$

The Model - Assumptions continues

- **Conditional Independence (CI):** State transition probability factorize as

$$\begin{aligned}\mathbb{P}\{i', \boldsymbol{\epsilon}'_{ig} \mid i, \boldsymbol{\epsilon}, a\} &= h(\boldsymbol{\epsilon} \mid i) \cdot \mathcal{F}_a(i' \mid i) \\ \mathbb{P}\{j', \boldsymbol{\epsilon}'_{jG} \mid j, \boldsymbol{\epsilon}, a\} &= h(\boldsymbol{\epsilon} \mid i) \cdot \mathcal{R}_a(j' \mid j).\end{aligned}$$

- $\mathcal{F}_a(i' \mid i)$ is the transition probability that a type i male g will next find himself single at age i' given his action a at age i.
- $\mathcal{R}_a(j' \mid j)$ is the transition probability that a type j female G will next find herself single at age j' given her action a.
- $\boldsymbol{\epsilon}$ are i.i.d. Type I Extreme Value random variables.
- full commitment, transferable utility setup.

The Model - Utility Functions

- If male g (or female G) chooses to marry an age j female (or i male),

$$\begin{aligned} v(a_{ig} = j, i, \boldsymbol{\epsilon}_{ig}) &= \boldsymbol{\alpha}_i(j) - \tau_{ij} + \epsilon_{ijg}, \text{ and} \\ w(a_{jG} = i, j, \boldsymbol{\epsilon}_{jG}) &= \boldsymbol{\gamma}_j(i) + \tau_{ij} + \epsilon_{ajG} \end{aligned}$$

$$\text{where } \boldsymbol{\alpha}_i(j) = \sum_{k=0}^{z_{ij}} (\beta S)^k \, \alpha_{ijk}, \qquad \text{and } \boldsymbol{\gamma}_j(i) = \sum_{k=0}^{z_{ij}} (\beta S)^k \, \gamma_{ijk}.$$

- α_{ijk} (or γ_{ijk}) be the $k'th$ period marital output accrued to a type i male (j female) when married to a type j female (or i male) today.

- If male g (or female G) chooses to remain single, then

$$v(a_{ig} = 0, i, \boldsymbol{\epsilon}_{ig}) = \alpha_{i0} + \epsilon_{i0g}, \quad \text{and} \quad w(a_{jG} = 0, j, \boldsymbol{\epsilon}_{jG}) = \gamma_{0j} + \epsilon_{0jG}$$

The Model - Convenient representation

- Rust's framework permits Value function to have convenient form,

$$\begin{aligned} V_\alpha(i, \boldsymbol{\epsilon}_{ig}) &= \max_{a \in \mathcal{D}} \{\tilde{v}_{ia} + \epsilon_{iag}\} \\ W_\gamma(j, \boldsymbol{\epsilon}_{jG}) &= \max_{a \in \mathcal{D}} \{\tilde{w}_{aj} + \epsilon_{ajG}\} \end{aligned}$$

- where the mean components, $\tilde{v}_{ij}$ and $\tilde{w}_{ij}$ are also referred to as the *choice specific value functions* for type i males and j females respectively.

$$\begin{aligned} \tilde{w}_{ij} &= (\boldsymbol{\gamma}_j(i) + \tau_{ij})\mathbb{I}(i \neq 0) + \gamma_{0j}\mathbb{I}(i = 0) + \sum_{j'} \mathcal{R}_i(j' \mid j) \cdot \boldsymbol{W}_{j'} \\ \tilde{v}_{ij} &= (\boldsymbol{\alpha}_i(j) - \tau_{ij})\mathbb{I}(j \neq 0) + \alpha_{i0}\mathbb{I}(j = 0) + \sum_{i'} \mathcal{F}_j(i' \mid i) \cdot \boldsymbol{V}_{i'}. \end{aligned}$$

- $\boldsymbol{V}_i$ and $\boldsymbol{W}_j$ are the integrated value function (value function where the unobservable state is integrated out)

$$\boldsymbol{V}_i = \int V_{\boldsymbol{\alpha}}(i, \boldsymbol{\epsilon}_g)\; dH(\boldsymbol{\epsilon}_g), \qquad \boldsymbol{W}_j = \int W_{\boldsymbol{\gamma}}(j, \boldsymbol{\epsilon}_G)\; dH(\boldsymbol{\epsilon}_G).$$

The Model - Choice Probabilities

- Define the *conditional choice probability* $\mathcal{P}_{ij}$ for males and $\mathcal{Q}_{ij}$ for females:

$$\begin{aligned}
\mathcal{P}_{ij} &= \int \mathbb{I}\{j = \arg\max_{a\in\mathcal{D}} (\tilde{v}_{ia} + \epsilon_{iag})\} h(d\boldsymbol{\epsilon}), \\
\mathcal{Q}_{ij} &= \int \mathbb{I}\{i = \arg\max_{a\in\mathcal{D}} (\tilde{w}_{aj} + \epsilon_{ajG})\} h(d\boldsymbol{\epsilon}).
\end{aligned}$$

- The probabilities have the familiar multinomial logit form,

$$\mathcal{P}_{ij} = \frac{\exp(\tilde{v}_{ij} - \tilde{v}_{i0})}{1 + \sum_{r=1}^{Z} \exp(\tilde{v}_{ir} - \tilde{v}_{i0})}, \qquad \mathcal{Q}_{ij} = \frac{\exp(\tilde{w}_{ij} - \tilde{w}_{i0})}{1 + \sum_{r=1}^{Z} \exp(\tilde{w}_{rj} - \tilde{w}_{0j})}.$$

The Model - Quasi Demand and Supply

- Log-odds ratios delivers a system of $(Z \times Z)$ quasi-demand and quasi-supply equations respectively.

$$\begin{aligned}
\ln \mathcal{P}_{ij} - \sum_{k=0}^{z_{ij}} (\beta S)^k \ln \mathcal{P}_{i+k,0} &= \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_i(0) - \tau_{ij} - \kappa \\
\ln \mathcal{Q}_{ij} - \sum_{k=0}^{z_{ij}} (\beta S)^k \ln \mathcal{Q}_{0,j+k} &= \boldsymbol{\gamma}_j(i) - \boldsymbol{\gamma}_j(0) + \tau_{ij} - \kappa.
\end{aligned}$$

where $\kappa = c\beta S(1-(\beta S)^{z_{ij}})/(1-\beta S)$, ($c$ is the Euler's constant)

$$\boldsymbol{\alpha}_i(j) = \sum_{k=0}^{z_{ij}} (\beta S)^k \alpha_{ijk}, \qquad \boldsymbol{\gamma}_j(i) = \sum_{k=0}^{z_{ij}} (\beta S)^k \gamma_{ijk},$$

$$\boldsymbol{\alpha}_i(0) = \sum_{k=0}^{z_{ij}} (\beta S)^k \alpha_{i+k,0}, \qquad \text{and } \boldsymbol{\gamma}_j(0) = \sum_{k=0}^{z_{ij}} (\beta S)^k \gamma_{0,j+k}.$$

The Model - Equilibrium

A marriage market equilibrium consists of a vector of males, $\boldsymbol{m}$ and females, $\boldsymbol{f}$ across individual type, the vector of marriage $\boldsymbol{\mu}$, and the vector of transfers, $\boldsymbol{\tau}$ such that the number of i type men who want to marry j type spouses exactly equals the number of j type women who agree to marry type i men for all combinations of (i, j). That is, for each of the $(Z \times Z)$ sub-markets,

$$m_i \mathcal{P}_{ij} = f_j \mathcal{Q}_{ij} = \mu_{ij}$$

$\square$

The Model - Dynamic Marriage Matching Function

- Let p_{ij} and q_{ij} denote the maximum likelihood estimators of $\mathcal{P}_{ij}$ and $\mathcal{Q}_{ij}$, that is, $p_{ij} = \mu_{ij}/m_i$ and $q_{ij} = \mu_{ij}/f_j$.

$$\mu_{ij} = \Pi_{ij}\sqrt{m_i f_j}\prod_{k=0}^{z_{ij}}\left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k}f_{j+k}}\right)^{\frac{1}{2}(\beta S)^k}$$

where $\Pi_{ij} = \sum_{k=0}^{z_{ij}}(\beta S)^k\left[(\alpha_{ijk}+\gamma_{ijk}) - (\alpha_{i+k,0}+\gamma_{0,j+k})\right] - 2\kappa$

The Model - Dynamic Marriage Matching Function

- The dynamic marriage matching function also needs to satisfy the accounting constraints given by,

$$
\begin{aligned}
\mu_{0j} + \sum_{i=1}^{Z} \mu_{ij} &= f_j \ \forall \ j \\
\mu_{i0} + \sum_{j=1}^{Z} \mu_{ij} &= m_i \ \forall \ i \\
\mu_{0j}, \mu_{i0}, \mu_{ij} &\geq 0 \ \forall \ i, j
\end{aligned}
$$

Inverse Problem

- Given a matrix of preferences $\mathbf{\Pi}$, whose elements are non-negative and strictly positive population vectors, $\boldsymbol{m}$ and $\boldsymbol{f}$, does there exist a unique non-negative marital distribution $\boldsymbol{\mu}$ that is consistent with $\mathbf{\Pi}$, that satisfies *Dynamic Marriage Matching Function* and accounting constraints.

- Reformulate the model to an $I+J$ system with $I+J$ number of unmarrieds of each type, μ_{i0} and μ_{0j}, as unknowns. This reduced system is defined by

$$m_i - \mu_{i0} = \sum_{i=1}^{I} \Pi_{ij} \sqrt{m_i f_j} \prod_{k=0}^{z_{ij}} \left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k} f_{j+k}} \right)^{\frac{1}{2}(\beta S)^k} \tag{3}$$

$$f_j - \mu_{0j} = \sum_{j=1}^{J} \Pi_{ij} \sqrt{m_i f_j} \prod_{k=0}^{z_{ij}} \left(\frac{\mu_{i+k,0}\mu_{0,j+k}}{m_{i+k} f_{j+k}} \right)^{\frac{1}{2}(\beta S)^k} \tag{4}$$

Existence and Uniqueness

- **Existence:** Generally the matching model with transferable utilities is equivalent to an optimal transportation (Monge-Kantorovich) linear programming problem.
- Optimal assignment in (Monge-Kantorovich) linear programming problem correspond to stable matching - optimal assignment shown to exist under mild conditions.
- See Chiappori, McCann and Nesheim (2009)
- **Uniqueness:** Linear programming models on compact convex feasible set have generically unique solutions. However for finite population, stable matching is generally not unique.

Empirical Application - Data

- Use model to describe changes in the gains to marriage in US from - 1970 to 1990
- Period of significant demographic and social changes: baby boomers, legalization of abortion, unilateral divorce, the pill, labor market changes, etc.
- Evaluate the importance of dynamics - compare model results with Choo and Siow (2006).
- Use Vital Statistics for marriages, μ_{ij} in 71/72, 81/82 and 91/92 from reporting states - individuals age between 16-75.
- Use 1970, 1980 1990 Census to get at unmarrieds, μ_{i0} and μ_{0j} (again matched on reporting states).

Empirical Application - Data Summary

Table 1a

A: US Census Data

	1970	1980	1990
Number of Available Males, (mill.)	16.018	23.412	28.417
Percentage change		46.2	21.4
Number of Available Females, (mill.)	19.592	27.225	31.563
Percentage change		39.0	15.9
Average age of Available Males	30.4	29.6	31.7
Average age of Available Females	39.1	37.1	37.9

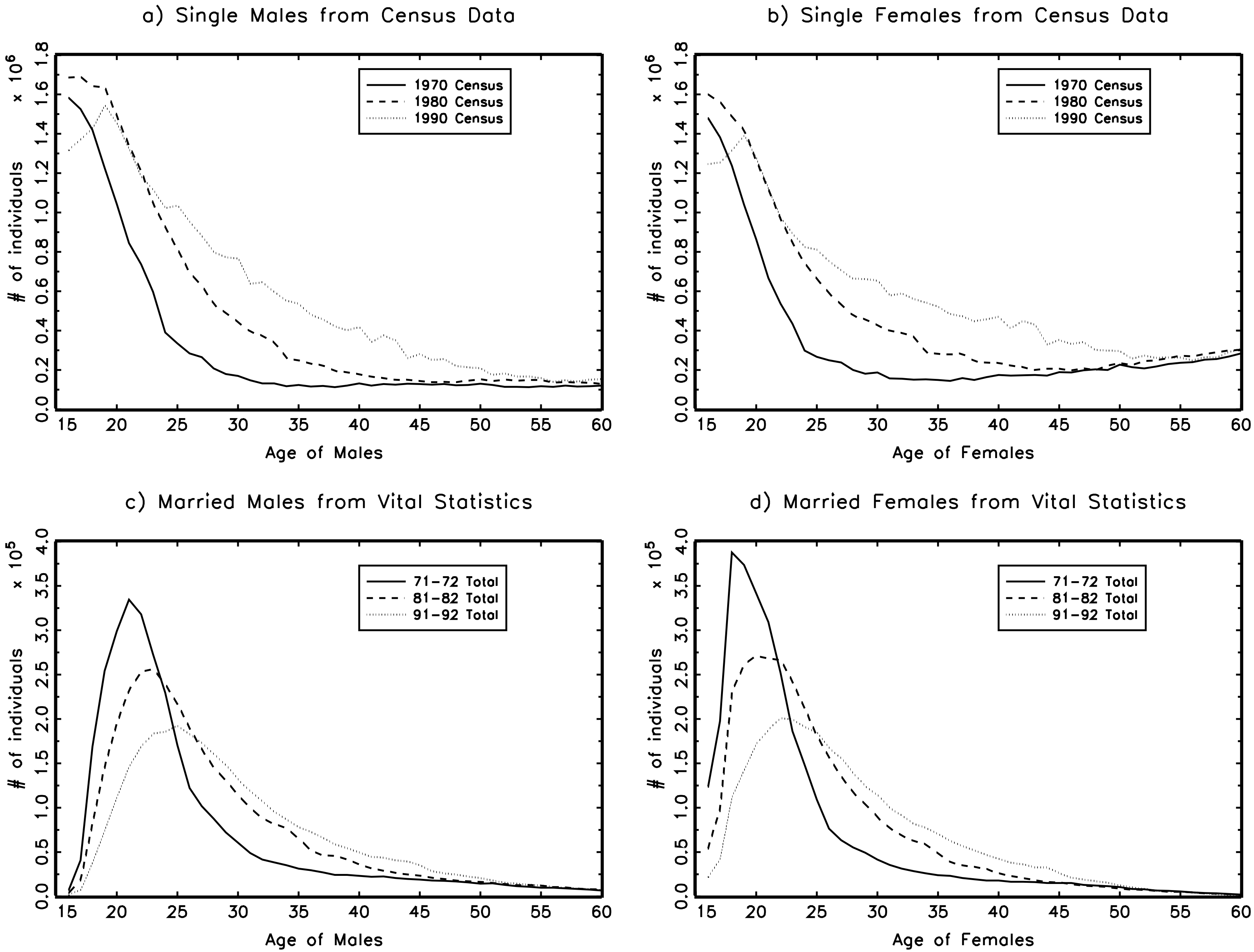
Empirical Application - Data Summary continues

Table 1b

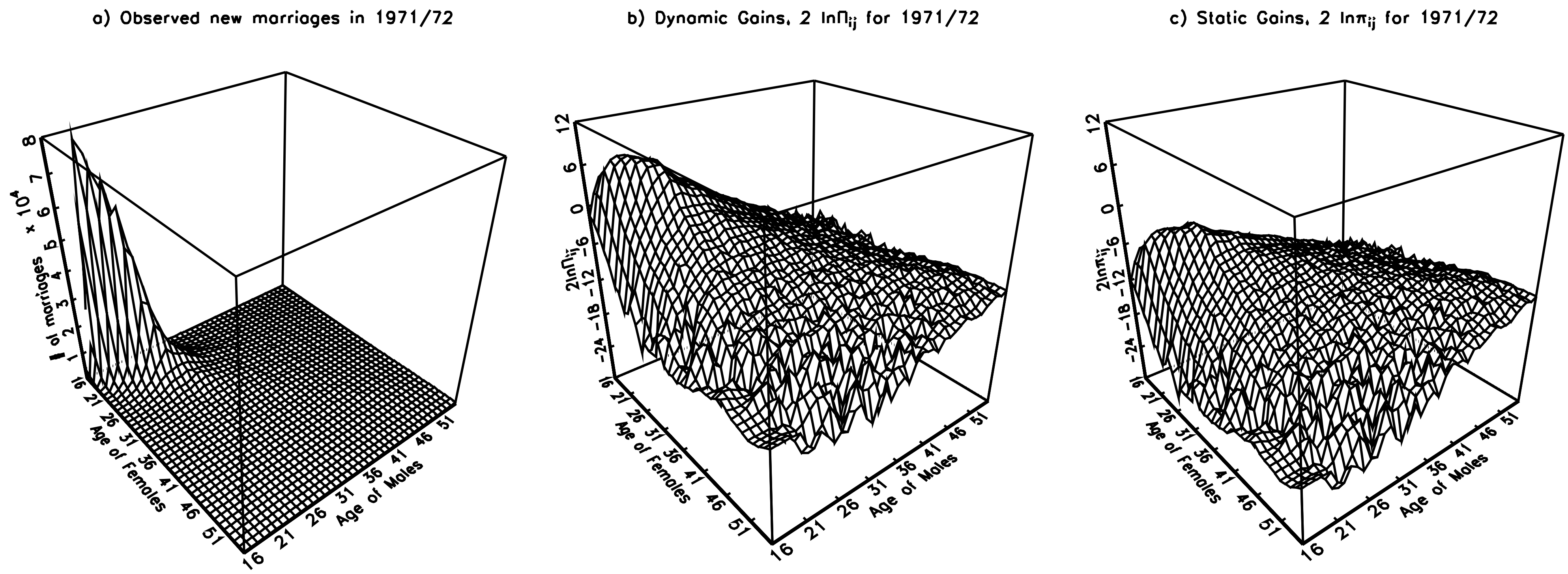
B: Vital Statistics Data

	1969-71	1979-81	1989-91
Average Number of marriages (mill.)	3.236	3.449	3.220
Percentage change		6.6	-7.11
Average age of Married Males	27.1	29.2	31.2
Average age of Married Females	24.5	26.4	28.9
Average couple age difference	2.6	2.7	2.3

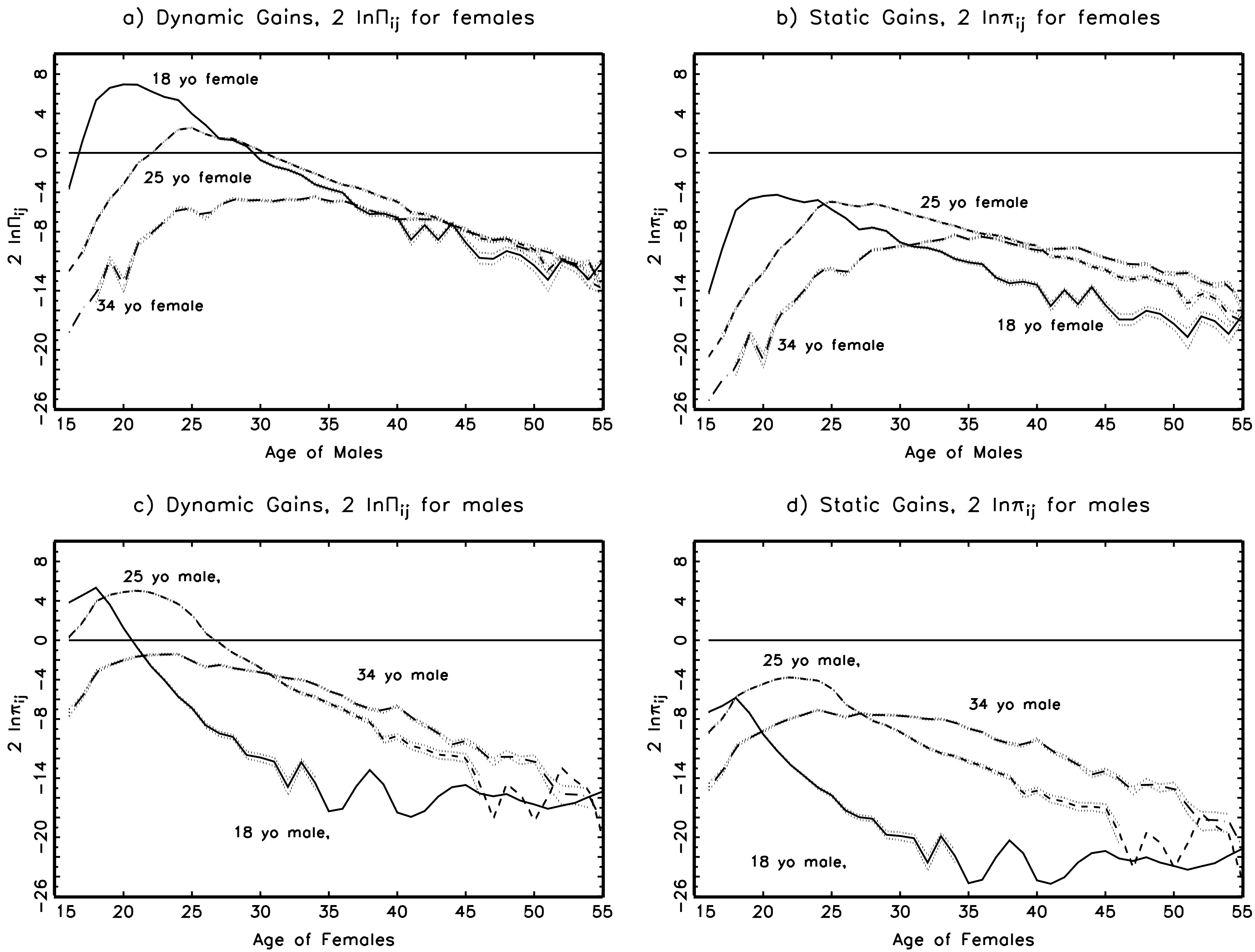
Plot of Singles and Married from 1970-1990



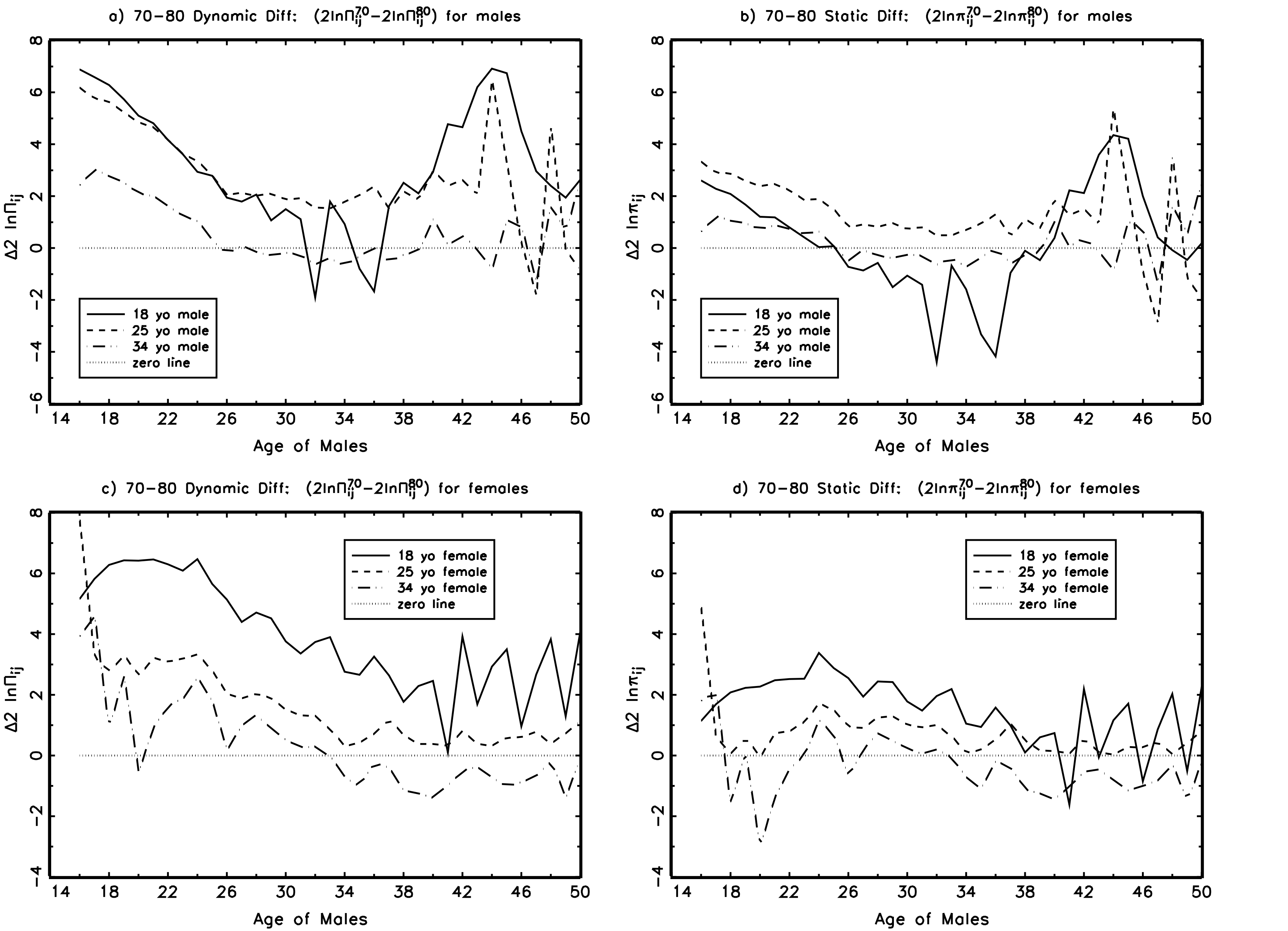
Comparing Dynamic and Static Gains for 71/72 μ_{ij}



Comparing Dynamic and Static Gains by gender for 71/72 μ_{ij}



Comparing Changes between 70-80 in Static and Dynamic Gains



Test for Model

- Rewrite quasi-demand and supply in terms of the maximum likelihood estimators p_{ij} and q_{ij}. That is,

$$\ln\left(p_{ij}\Big/\prod_{k=0}^{z_{ij}} p_{i+k,0}^{(\beta S)^k}\right) = \boldsymbol{\alpha}_i(j) - \boldsymbol{\alpha}_i(0) - \tau_{ij} - \kappa,$$

$$\ln\left(q_{ij}\Big/\prod_{k=0}^{z_{ij}} q_{i+k,0}^{(\beta S)^k}\right) = \boldsymbol{\gamma}_j(i) - \boldsymbol{\gamma}_j(0) + \tau_{ij} - \kappa.$$

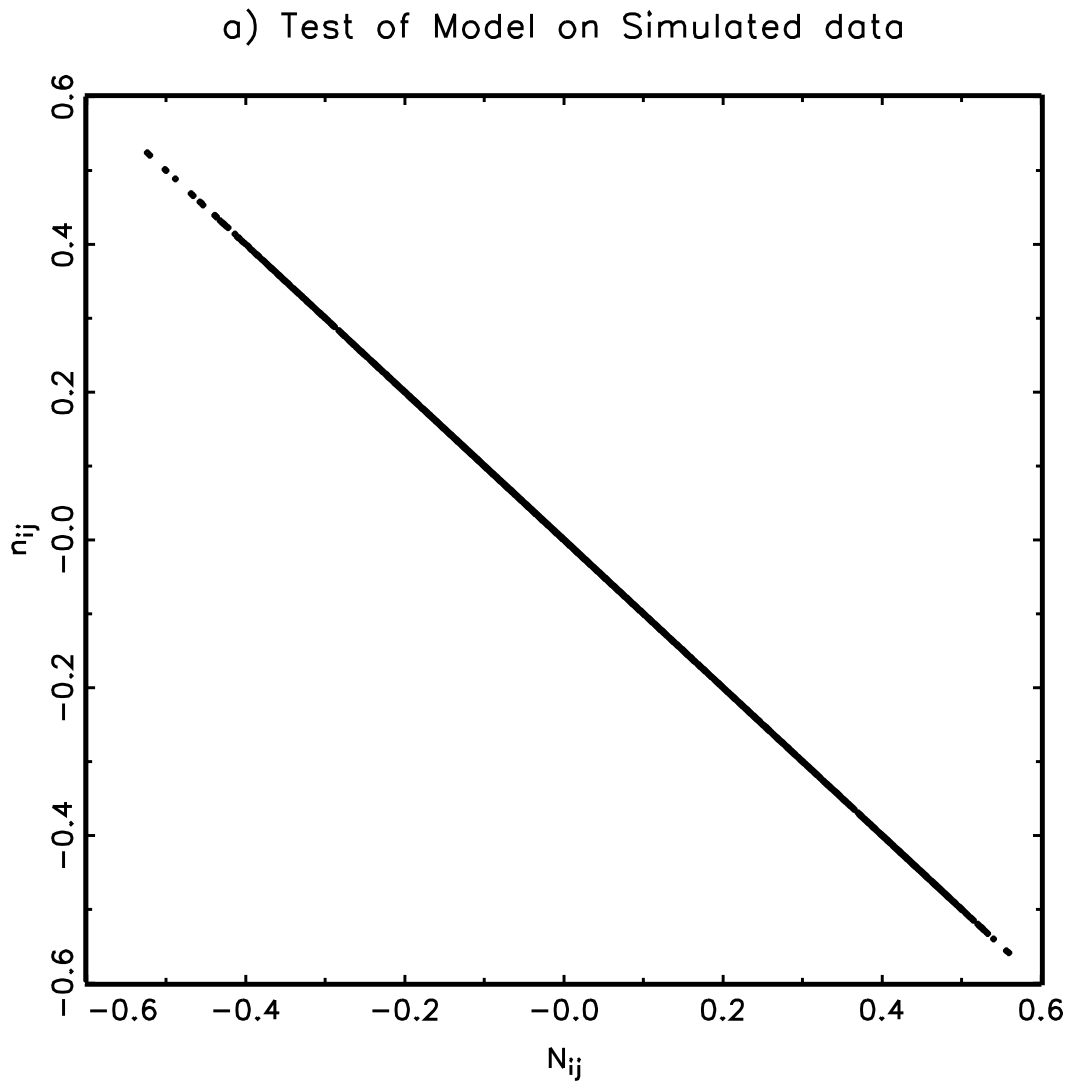
Let $n_{ij}(\boldsymbol{\mu}, \boldsymbol{m}, \boldsymbol{f}) = \ln\left(p_{ij}\Big/\prod_{k=0}^{z_{ij}} p_{i+k,0}^{(\beta S)^k}\right)$ and

$$\mathcal{N}_{ij}(\boldsymbol{\mu}, \boldsymbol{m}, \boldsymbol{f}) = \ln\left(q_{ij}\Big/\prod_{k=0}^{z_{ij}} q_{i+k,0}^{(\beta S)^k}\right)$$

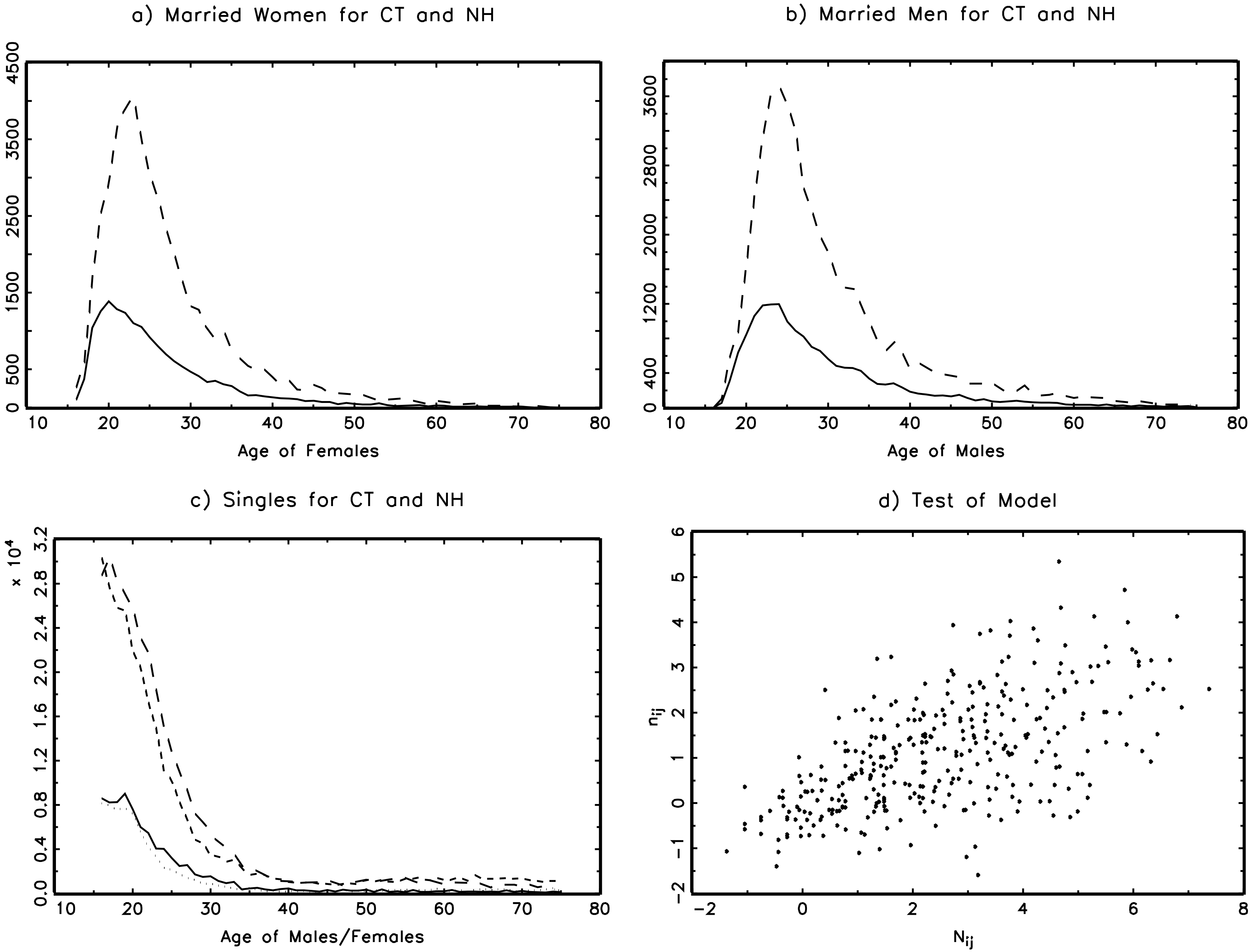
Test for Model continues

Proposition 2: Holding α_{ijk}, γ_{ijk}, and δ_{ijk} fixed for all (i, j, k), any changes in available men m_i or women f_j that leads to an increase in $n_{ij}(\boldsymbol{\mu}, \boldsymbol{m}, \boldsymbol{f})$ would also lead to a decrease in $\mathcal{N}_{ij}(\boldsymbol{\mu}, \boldsymbol{m}, \boldsymbol{f})$ and vice versa.

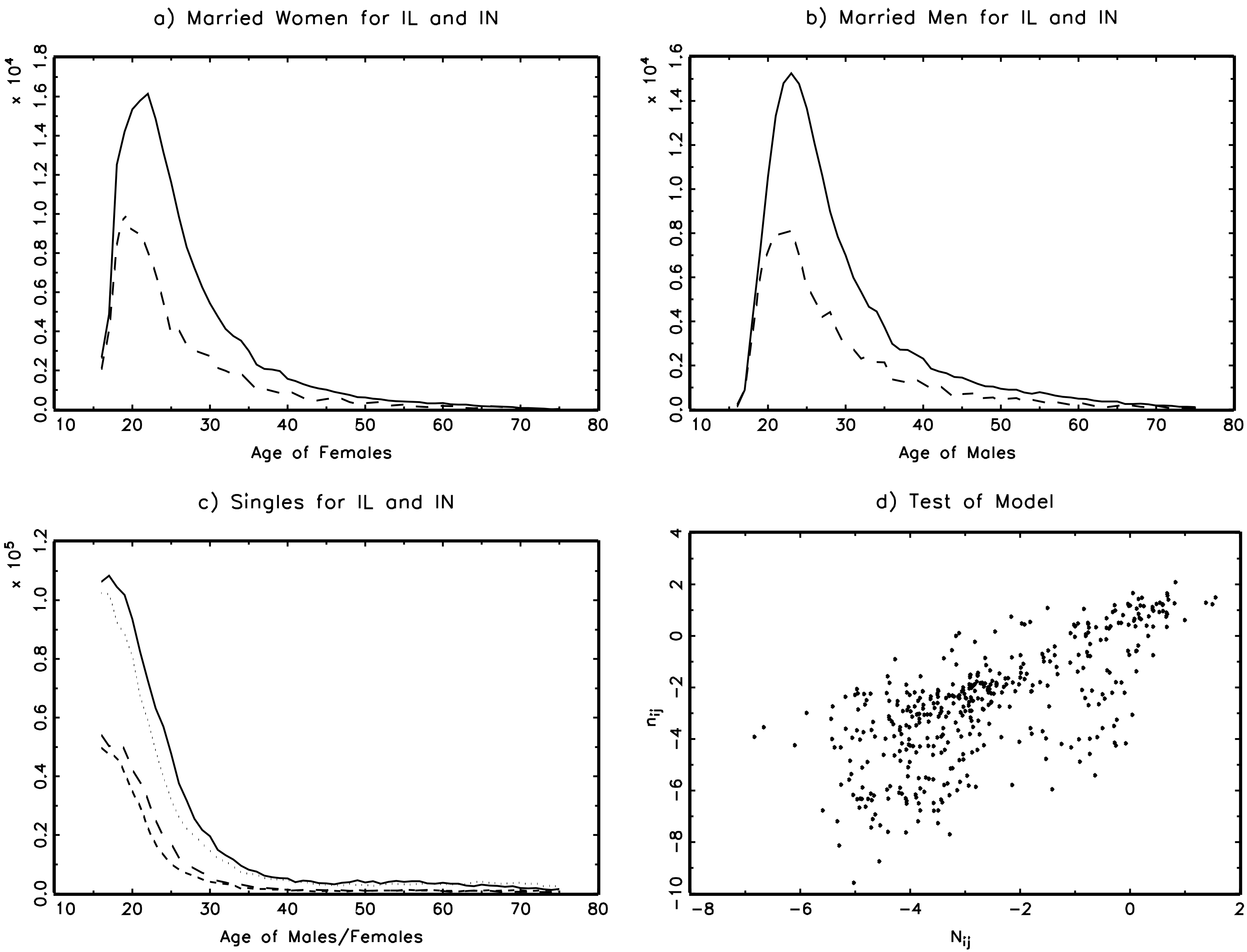
Plot of $\mathcal{N}_{ij}$ and n_{ij} on simulated data



Comparing CT and NH



Comparing IL and IN



Conclusion

- Proposed an tractable dynamic marriage matching model that maintains many of the convenient properties of the static Choo and Siow (2006) model.
- Demonstrate that the dynamic components to marital returns is large among the young.
- Also propose a test for the model.