Why People Migrate:
The Diverse Mechanisms of Mexico-U.S. Migration[1]

Filiz Garip

Harvard University

The Hidden Diversity of Migration

There are diverse mechanisms that lead individuals to migrate to the United States.

This diversity is captured in various migration theories developed in multiple disciplines.

But, it gets lost in empirical studies that describe a typical narrative for an average migrant.

Overview of Migration Theories

Neoclassical economics

Migration is an individual strategy to increase income.

New economics of labor migration

Migration is a household act to diversify risks to household income.

Cumulative causation

Migration creates social ties between origin and destination, which increase future movement.

A Gap Between Theory and Evidence

Each theory applies to a specific group of individuals under specific conditions.

The predictions of different migration theories are *not* mutually exclusive.

Using regression analysis, empirical studies neglect…

   …the *conditional* nature of each theory, and

   …the *complementarity* among different theories.

Bridging the Gap

Prior work considered the causal heterogeneity of migration using <u>regression models</u> with
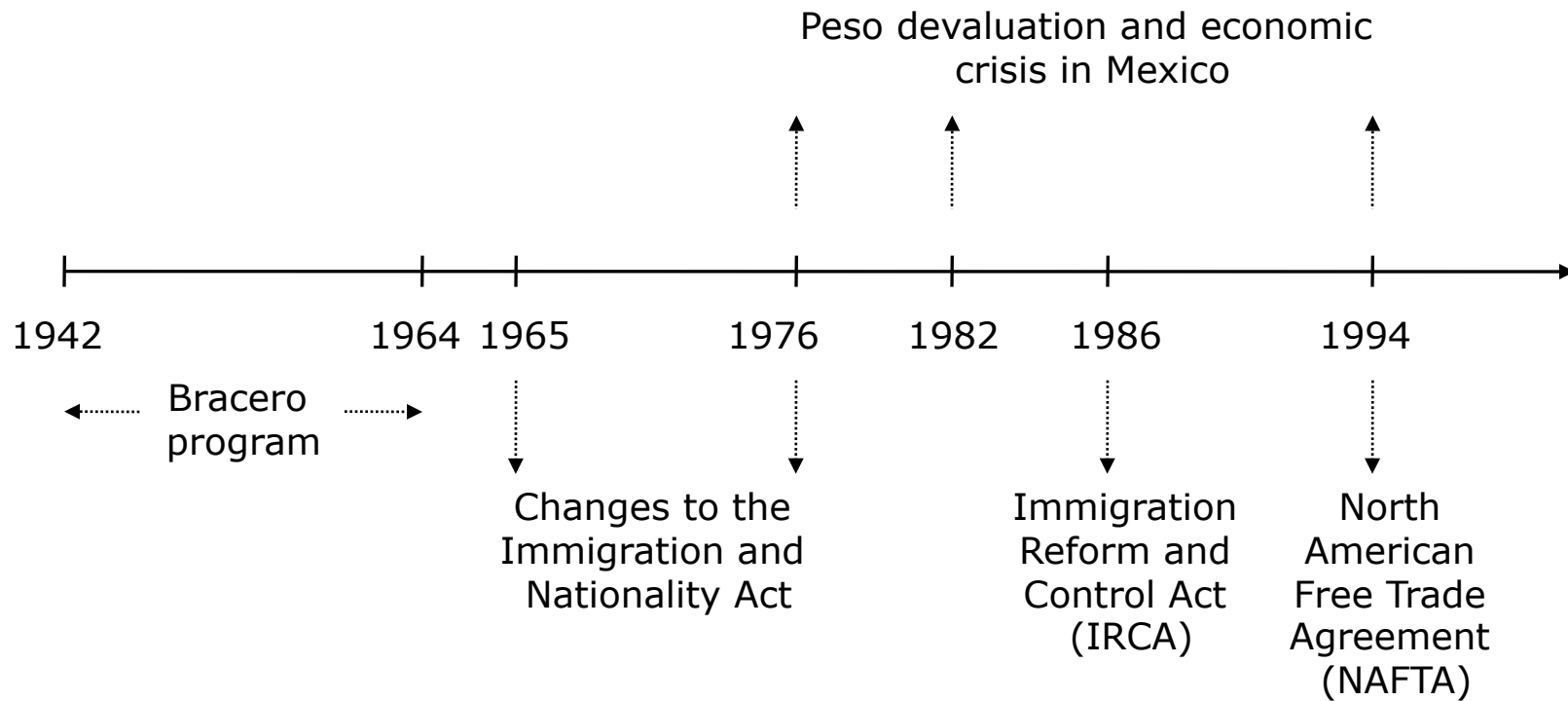
...split samples, or

...interaction terms.

As an alternative, this study uses <u>cluster analysis</u>…

…to identify the *distinct configurations of factors*

…that characterize different *migrant types*.

Migration from Mexico to the United States

Major milestones

Peso devaluation and economic
crisis in Mexico

1942   1964   1965   1976   1982   1986   1994

Bracero
program

Changes to the
Immigration and
Nationality Act

Immigration
Reform and
Control Act
(IRCA)

North
American
Free Trade
Agreement
(NAFTA)

Study Data

Mexican Migration Project (MMP) surveyed…

    …about 200 randomly-selected households

    …in 124 communities from 21 states

    …between 1987 and 2008.

Our sample contains…

    …17,049 migrants

    …observed during their first U.S. trip

    …between 1970 to 2000.

# A Brief Introduction to Cluster Analysis

Cluster analysis is a method for discovering groups with similar attributes in the data.

It is widely used in data-intensive fields like biology, physics or computer science, but not in the social sciences.

Different than regression analysis, which searches for associations to an outcome, cluster analysis identifies groups in data based on the variability in several attributes.

It is particularly useful when those attributes are 'clustered' around distinct configurations.

An Example in 2 Dimensions
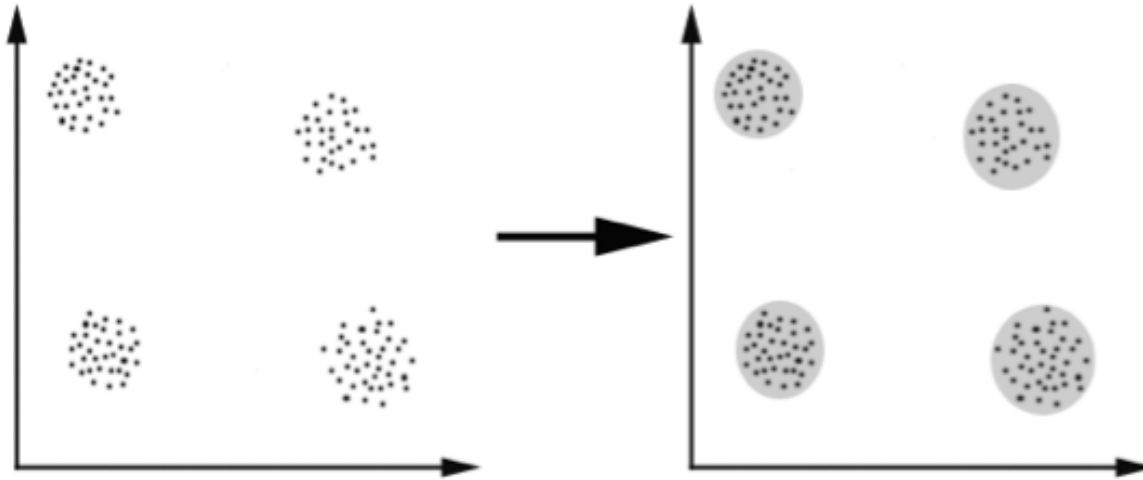
Steps in Cluster Analysis (I)

Choosing the Relevant Attributes

Similar to variable selection in regression analysis

Rely on theory or data to identify salient attributes

e.g., use prior empirical work on the MMP data

Attribute Means for Migrants and Non-migrants (I)

	Migrants	Non-migrants
Demographic characteristics		
Household head	**0.28**	0.12
Male	**0.72**	0.45
Years of education	**6.86**	5.97
Agricultural occupation	**0.20**	0.07
Manufacturing occupation	**0.31**	0.10
Service occupation	**0.22**	0.12
Unemployed	0.19	**0.64**
Household wealth		
Number of rooms in properties	**4.21**	3.61
Log of land value (in US$ in 2000)	**3.12**	2.15
Number of businesses	**0.41**	0.37

The higher mean value is shown in boldface, and differs significantly between migrants and non-migrants (p<0.05, two-tailed test).

Attribute Means for Migrants and Non-migrants (II)

	Migrants	Non-migrants
Migration experience		
Migrated in Mexico?	**0.22**	0.17
No. of U.S. legal residents in household	**0.48**	0.13
No. of U.S. migrants (non-residents) in household	**1.81**	0.72
Proportion ever migrated in community	**0.19**	0.13
Community characteristics		
Proportion in agriculture in community	**0.28**	0.24
Proportion self-employed in community	**0.32**	0.30
Proportion earning less than min. wage in community	**0.38**	0.34
Community in metropolitan area	0.41	**0.50**
N (persons)	17,049	107,838

The higher mean value is shown in boldface, and differs significantly between migrants and non-migrants (p<0.05, two-tailed test).

Steps in Cluster Analysis (I)

Choosing the Relevant Attributes

Similar to variable selection in regression analysis

Rely on theory or data to identify salient attributes

e.g., use prior empirical work on the MMP data

Scale the attributes to avoid arbitrary weighting

e.g., standardize the range or the variance

Steps in Cluster Analysis (II)

Choosing the Algorithm

Divides the data into groups based on the selected attributes

Output includes…

…a *cluster membership* for each case, and

…a *centroid* for each cluster that represents the mean of the cases in that cluster

A popular choice, also used here, is the K-means method.

Alternatives include hierarchical or model-based methods.

Steps in Cluster Analysis (III)

Choosing a Similarity Measure

Criterion to determine how 'close' cases are to one another in the attribute space

More consequential than choosing the algorithm

Decision depends on the nature of the data and the substance of the question

This study uses the city block distance.

$$d_{ij} = \left(\sum_{l=1}^{p} \left| x_{il} - x_{jl} \right| \right)$$

Steps in Cluster Analysis (IV)

Choosing the Number of Clusters

K-means method requires the researcher to supply K, i.e., the number of clusters

The algorithm finds K clusters *even when no such structure exists in data*

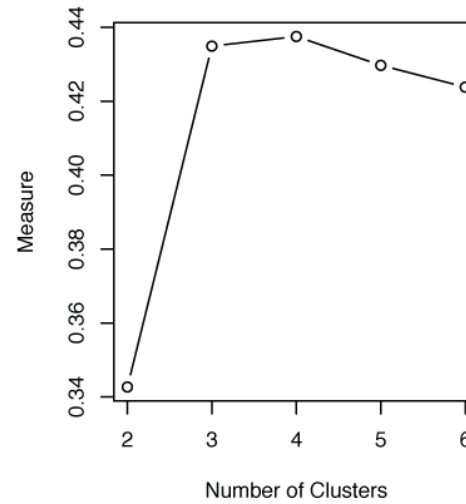To avoid artificial partitions, use cluster validation measures to choose K

~ similar to model selection in regression analysis
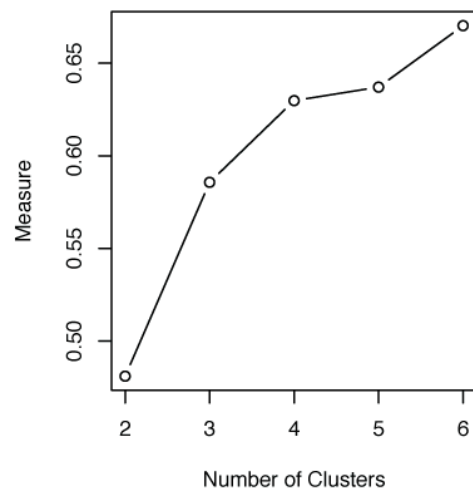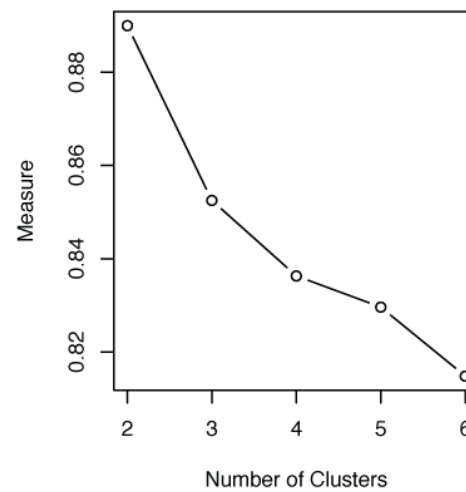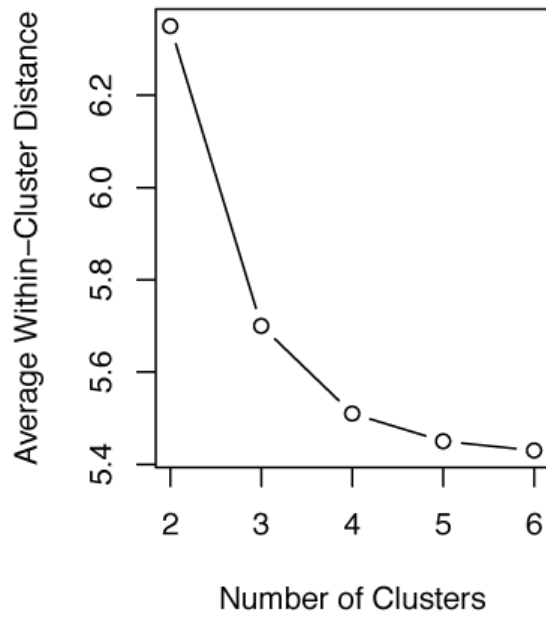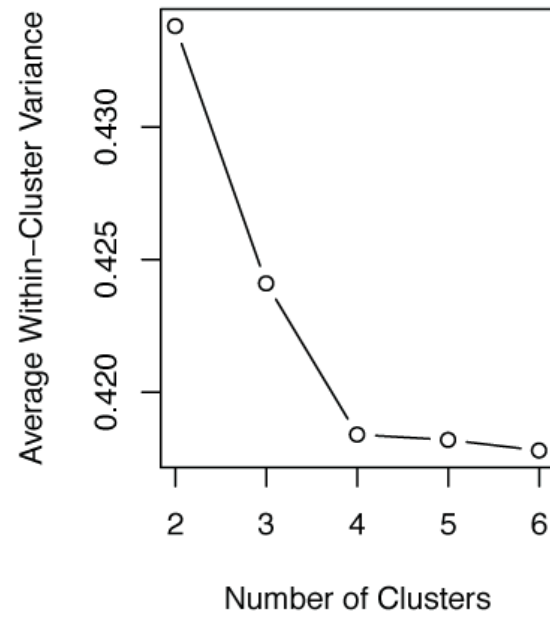
Cluster Validation Measures
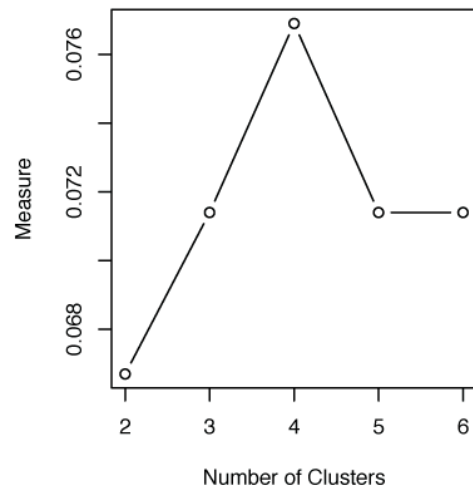
Cluster Stability Measures

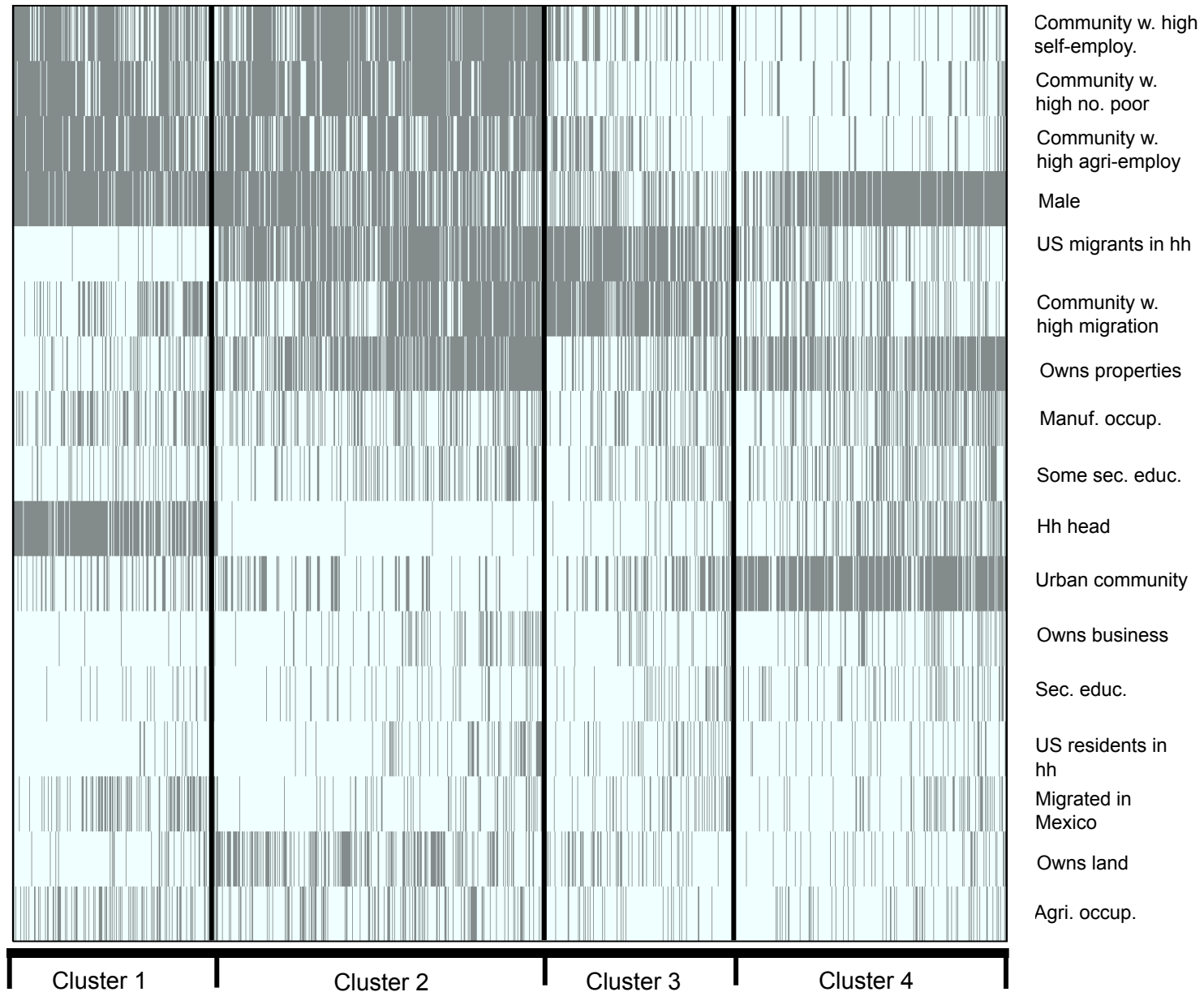# Steps in Cluster Analysis (V)

Assessing the Validity of Results

Draw a cluster heat map

Used in genetics to visualize gene expressions across samples

In our case, shows the distribution of attributes across migrants in the four clusters

Determine if the clusters are well-separated

Heat Map of Migrant Attributes by Cluster

Community w. high self-employ.
Community w. high no. poor
Community w. high agri-employ
Male
US migrants in hh
Community w. high migration
Owns properties
Manuf. occup.
Some sec. educ.
Hh head
Urban community
Owns business
Sec. educ.
US residents in hh
Migrated in Mexico
Owns land
Agri. occup.

Cluster 1 — Cluster 2 — Cluster 3 — Cluster 4

Migrant Attributes By Cluster (I)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Demographic characteristics				
Household head	**0.83**	0.02	0.07	0.32
Male	**0.90**	0.73	0.38	0.80
Some secondary education	0.17	0.24	0.24	**0.30**
Complete secondary education	0.08	0.09	0.12	**0.17**
Agricultural occupation	**0.31**	0.24	0.11	0.13
Manufacturing occupation	0.31	0.29	0.23	**0.39**
Household wealth				
Own properties	0.19	**0.76**	0.33	0.67
Own land	0.11	**0.38**	0.20	0.12
Own business	0.05	**0.16**	0.11	0.14

The highest mean value is shown in boldface, and differs significantly from the value closest to it ($p < 0.05$, two-tailed test) for all attributes.

Migrant Attributes By Cluster (II)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Migration experience				
Migrated in Mexico?	**0.33**	0.12	0.14	0.18
Any U.S. legal residents in household	0.04	0.14	**0.17**	0.09
Any U.S. migrants (non-residents) in household	0.04	0.80	**0.81**	0.33
Community with high migration prevalence	0.34	0.60	**0.79**	0.29
Community characteristics				
Community with high agriculture employment	**0.82**	0.74	0.27	0.12
Community with high self employment	0.69	**0.85**	0.27	0.10
Community with high no. of low earners	0.79	**0.83**	0.15	0.13
Community in metropolitan area	0.26	0.20	0.34	**0.81**
N (persons)	3,522	5,569	3,271	4,687
(% of total number of migrants)	21	33	19	27

The highest mean value is shown in boldface, and differs significantly from the value closest to it ($p<0.05$, two-tailed test) for all attributes but one (any U.S. migrants in the household).

Ideal Types by Cluster

A typical migrant in **Cluster 1**...

> ...is a male household head

> ...has no education

> ...owns no assets

> ...lives in a poor rural community

> ...migrates to increase income

> ...personifies neoclassical economics

> ...labeled an **income maximizer**

Migrant Attributes By Cluster (I)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Demographic characteristics				
Household head	**0.83**	0.02	0.07	0.32
Male	**0.90**	0.73	0.38	0.80
Some secondary education	0.17	0.24	0.24	**0.30**
Complete secondary education	0.08	0.09	0.12	**0.17**
Agricultural occupation	**0.31**	0.24	0.11	0.13
Manufacturing occupation	0.31	0.29	0.23	**0.39**
Household wealth				
Own properties	0.19	**0.76**	0.33	0.67
Own land	0.11	**0.38**	0.20	0.12
Own business	0.05	**0.16**	0.11	0.14

The highest mean value is shown in boldface, and differs significantly from the value closest to it ($p < 0.05$, two-tailed test) for all attributes.

Migrant Attributes By Cluster (II)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Migration experience				
Migrated in Mexico?	**0.33**	0.12	0.14	0.18
Any U.S. legal residents in household	0.04	0.14	**0.17**	0.09
Any U.S. migrants (non-residents) in household	0.04	0.80	**0.81**	0.33
Community with high migration prevalence	0.34	0.60	**0.79**	0.29
Community characteristics				
Community with high agriculture employment	**0.82**	0.74	0.27	0.12
Community with high self employment	0.69	**0.85**	0.27	0.10
Community with high no. of low earners	0.79	**0.83**	0.15	0.13
Community in metropolitan area	0.26	0.20	0.34	**0.81**
N (persons)	3,522	5,569	3,271	4,687
(% of total number of migrants)	21	33	19	27

The highest mean value is shown in boldface, and differs significantly from the value closest to it ($p<0.05$, two-tailed test) for all attributes but one (any U.S. migrants in the household).

25

Ideal Types by Cluster

A typical migrant in **Cluster 2**…

> …is a son in the household

> …has primary education

> …owns many assets

> …lives in a poor rural community

> …migrates to diversify risks to household income

> …acts in line with the new economics theory

> …labeled a **risk diversifier**

Migrant Attributes By Cluster (I)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Demographic characteristics				
Household head	**0.83**	0.02	0.07	0.32
Male	**0.90**	0.73	0.38	0.80
Some secondary education	0.17	0.24	0.24	**0.30**
Complete secondary education	0.08	0.09	0.12	**0.17**
Agricultural occupation	**0.31**	0.24	0.11	0.13
Manufacturing occupation	0.31	0.29	0.23	**0.39**
Household wealth				
Own properties	0.19	**0.76**	0.33	0.67
Own land	0.11	**0.38**	0.20	0.12
Own business	0.05	**0.16**	0.11	0.14

The highest mean value is shown in boldface, and differs significantly from the value closest to it ($p<0.05$, two-tailed test) for all attributes.

Migrant Attributes By Cluster (II)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Migration experience				
Migrated in Mexico?	**0.33**	0.12	0.14	0.18
Any U.S. legal residents in household	0.04	0.14	**0.17**	0.09
Any U.S. migrants (non-residents) in household	0.04	0.80	**0.81**	0.33
Community with high migration prevalence	0.34	0.60	**0.79**	0.29
Community characteristics				
Community with high agriculture employment	**0.82**	0.74	0.27	0.12
Community with high self employment	0.69	**0.85**	0.27	0.10
Community with high no. of low earners	0.79	**0.83**	0.15	0.13
Community in metropolitan area	0.26	0.20	0.34	**0.81**
N (persons)	3,522	5,569	3,271	4,687
(% of total number of migrants)	21	33	19	27

The highest mean value is shown in boldface, and differs significantly from the value closest to it ($p < 0.05$, two-tailed test) for all attributes but one (any U.S. migrants in the household).

28

Ideal Types by Cluster

A typical migrant in **Cluster 3**...

   ...is a daughter in the household

   ...is unemployed

   ...is connected to other U.S. migrants

   ...migrates to join her family

   ...personifies cumulative causation theory

   ...labeled a **network migrant**

Migrant Attributes By Cluster (I)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Demographic characteristics				
Household head	**0.83**	0.02	0.07	0.32
Male	**0.90**	0.73	0.38	0.80
Some secondary education	0.17	0.24	0.24	**0.30**
Complete secondary education	0.08	0.09	0.12	**0.17**
Agricultural occupation	**0.31**	0.24	0.11	0.13
Manufacturing occupation	0.31	0.29	0.23	**0.39**
Household wealth				
Own properties	0.19	**0.76**	0.33	0.67
Own land	0.11	**0.38**	0.20	0.12
Own business	0.05	**0.16**	0.11	0.14

The highest mean value is shown in boldface, and differs significantly from the value closest to it ($p < 0.05$, two-tailed test) for all attributes.

Migrant Attributes By Cluster (II)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Migration experience				
Migrated in Mexico?	**0.33**	0.12	0.14	0.18
Any U.S. legal residents in household	0.04	0.14	**0.17**	0.09
Any U.S. migrants (non-residents) in household	0.04	0.80	**0.81**	0.33
Community with high migration prevalence	0.34	0.60	**0.79**	0.29
Community characteristics				
Community with high agriculture employment	**0.82**	0.74	0.27	0.12
Community with high self employment	0.69	**0.85**	0.27	0.10
Community with high no. of low earners	0.79	**0.83**	0.15	0.13
Community in metropolitan area	0.26	0.20	0.34	**0.81**
N (persons)	3,522	5,569	3,271	4,687
(% of total number of migrants)	21	33	19	27

The highest mean value is shown in boldface, and differs significantly from the value closest to it ($p < 0.05$, two-tailed test) for all attributes but one (any U.S. migrants in the household).

31

Ideal Types by Cluster

A typical migrant in **Cluster 4**...

>...is son in the household

>...has some secondary education

>...owns property

>...lives in an urban community

>...is not anticipated in any migration theory

>...labeled an **urban migrant**

Trends in the Distribution of Migrants Across Clusters

Income Maximizers

Trends in the Distribution of Migrants Across Clusters
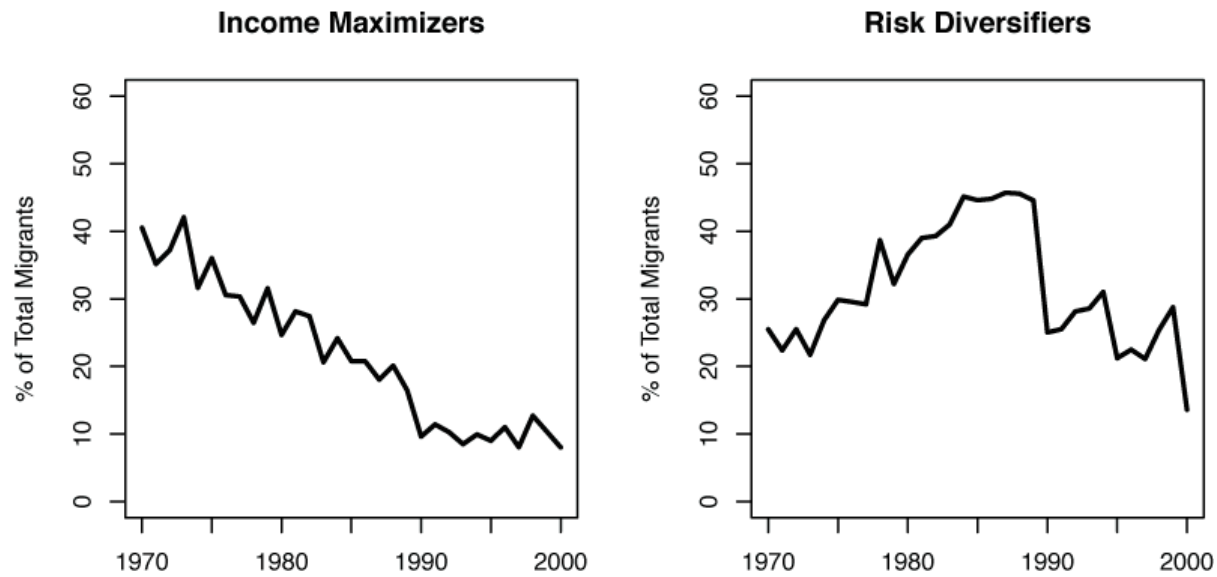
Trends in the Distribution of Migrants Across Clusters
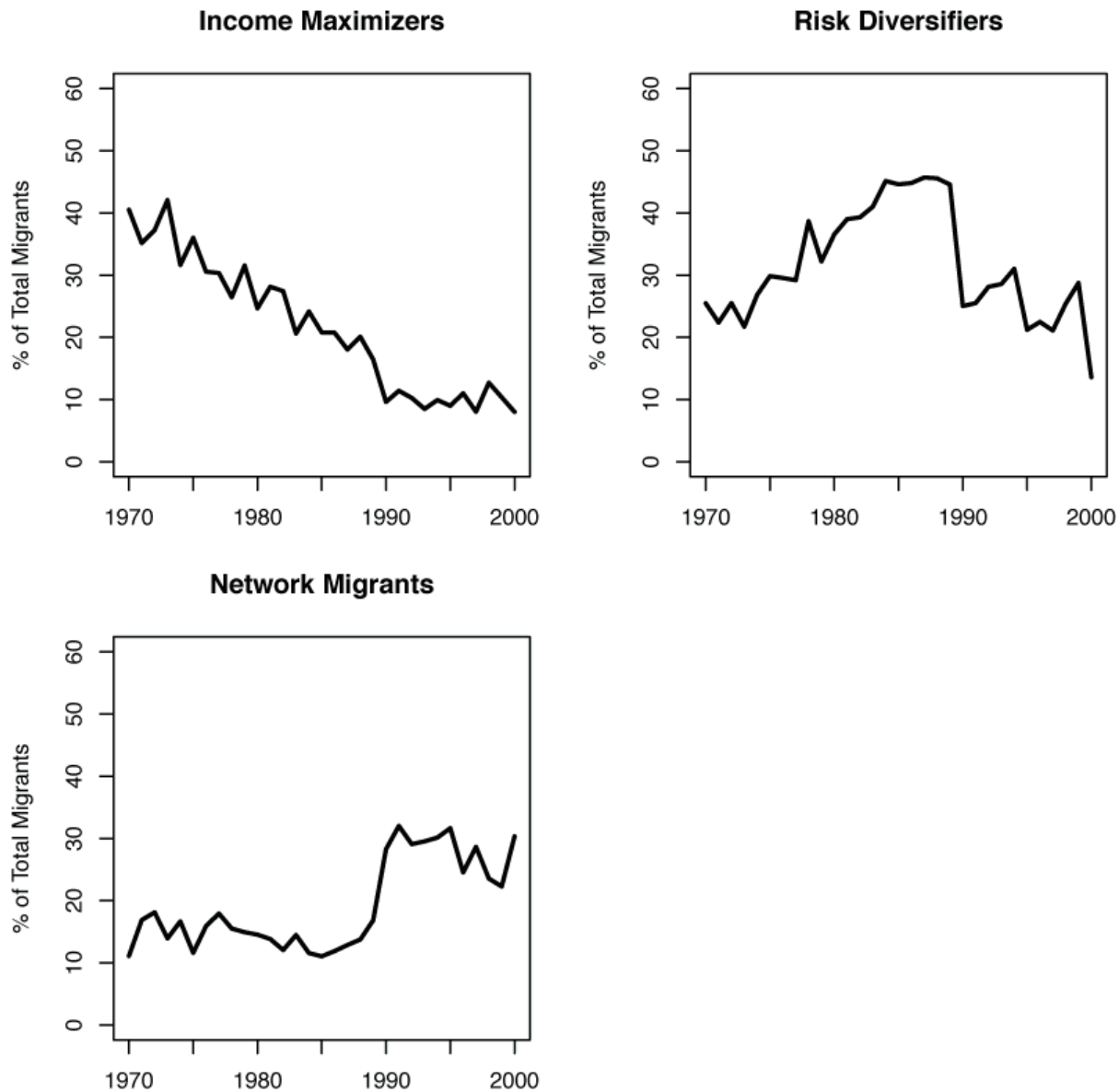
Trends in the Distribution of Migrants Across Clusters
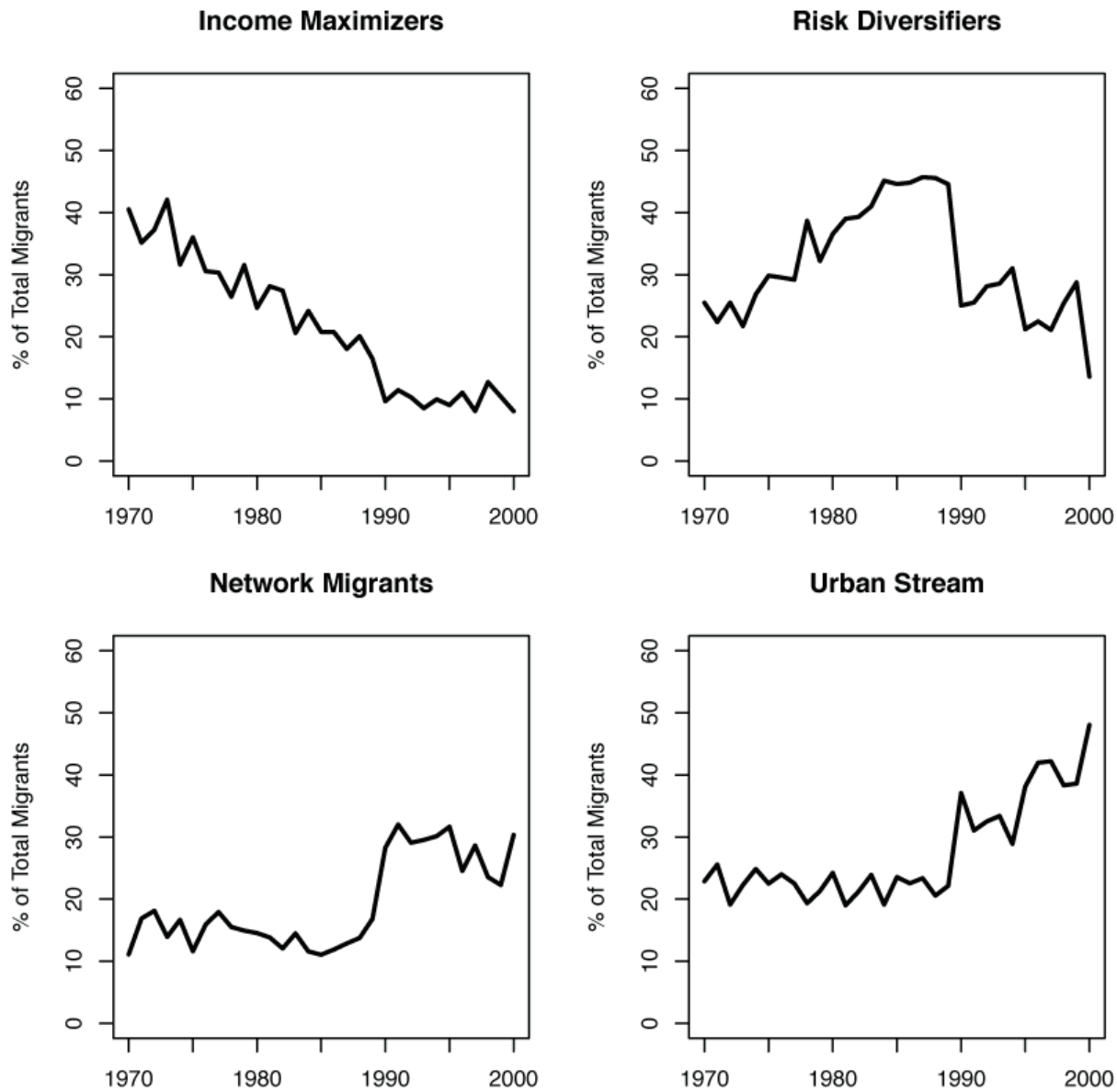
The Economic and Political Context (1960-1970)

Mexico…

Decline in agricultural productivity

Shortage of job opportunities

Worsening of living standards for rural families

United States…

End of the Bracero program

But farm work still defined as 'immigrant' work

Trends in the Economic and Political Context
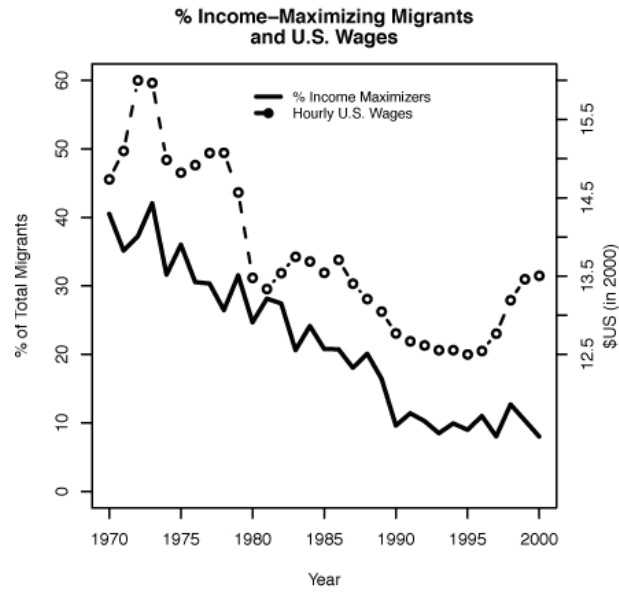
The Economic and Political Context (1970-1980)

Mexico…

   Peso devaluations in 1976 and 1982

   Increase in inflation and interest rates

   Reduction in agricultural subsidies

   Setbacks for middle-income rural families

Trends in the Economic and Political Context
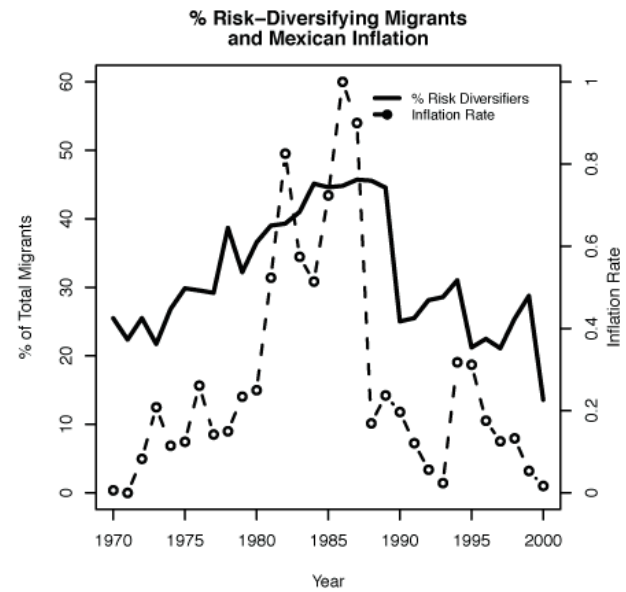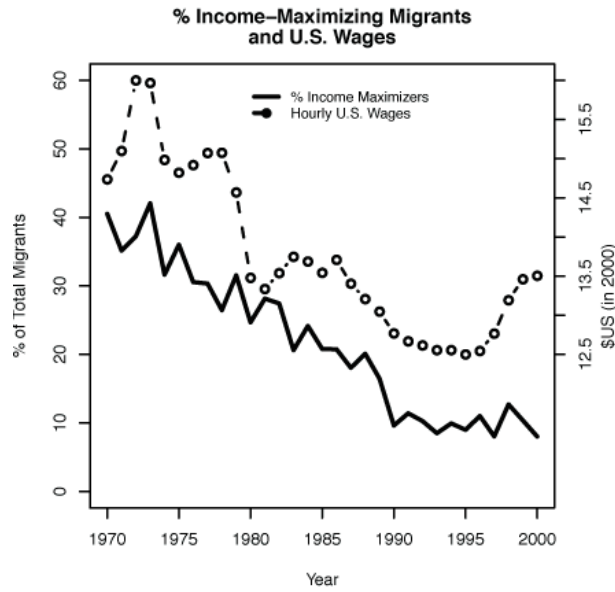
The Economic and Political Context (1980-2000)

United States…

Political backlash against undocumented migration

Passage of Immigration Reform and Control Act (IRCA) in 1986

…increased border enforcement

…increased sanctions on employers

…legalized 2.3 million Mexican migrants

Signs the Generalized Agreement on Tariffs and Trade (GATT) with Mexico in 1986

Trends in the Economic and Political Context
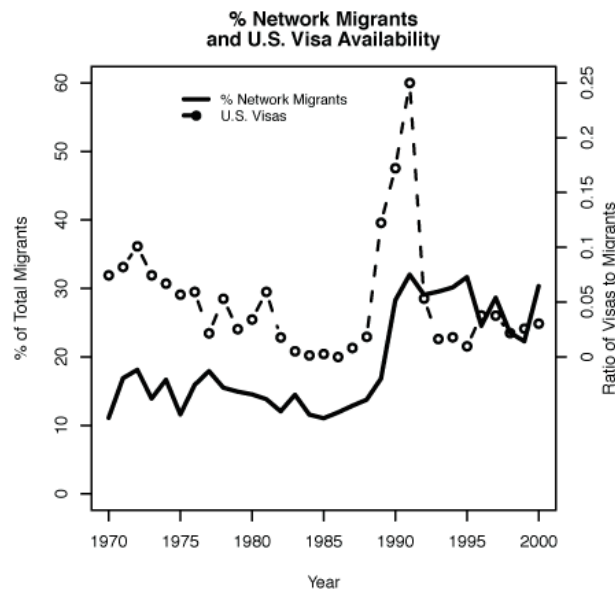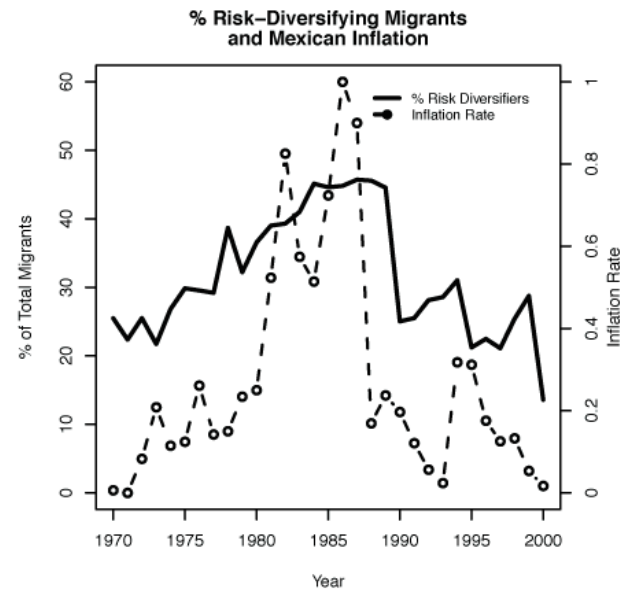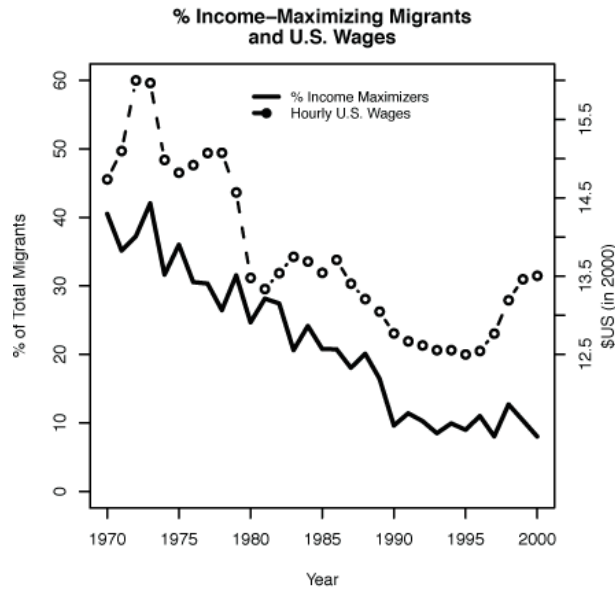
The Economic and Political Context (1990-2000)

Mexico…

Peso devaluation in 1994

Sharp increase in unemployment rate

United States…

Joins the North American Free Trade Agreement (NAFTA) with Mexico in 1994

Trends in the Economic and Political Context
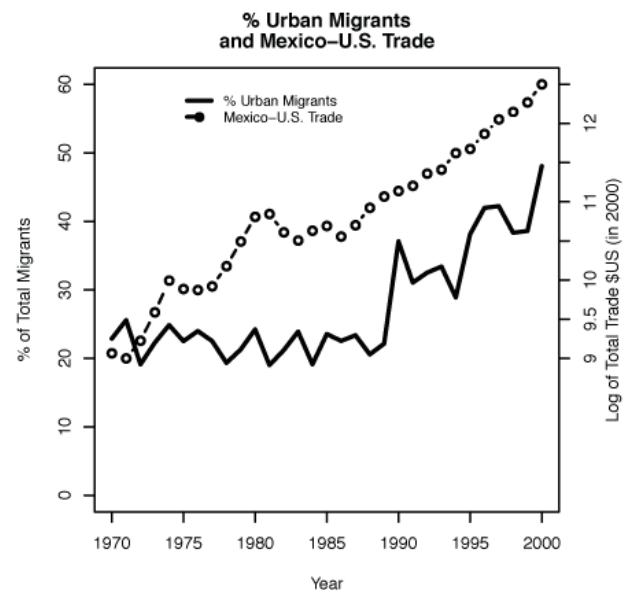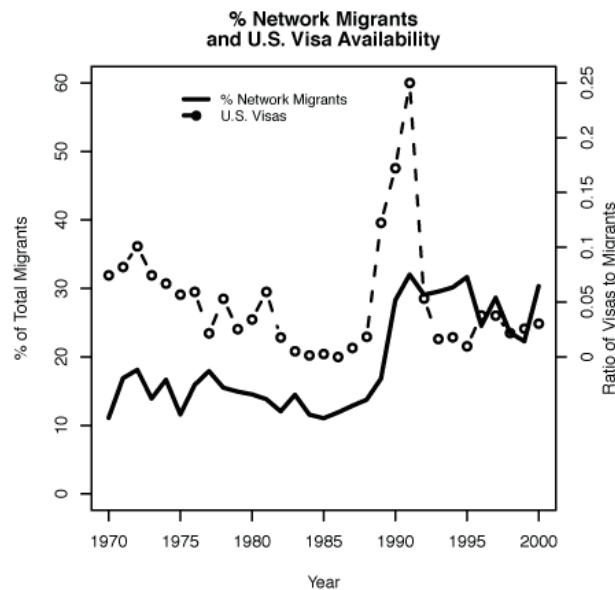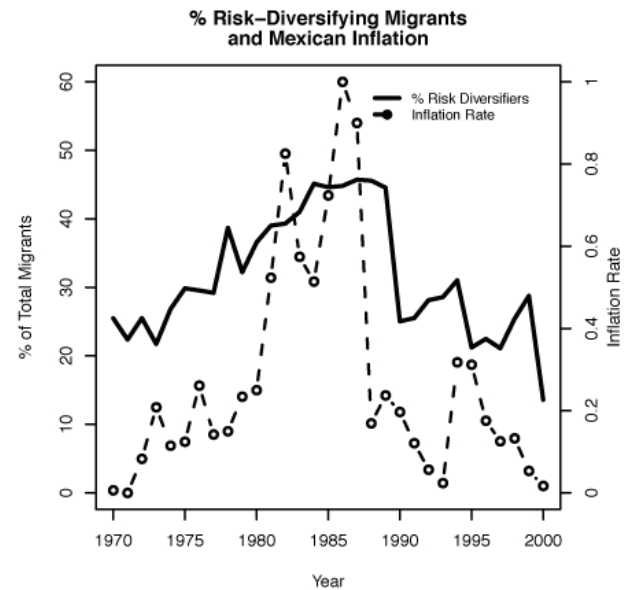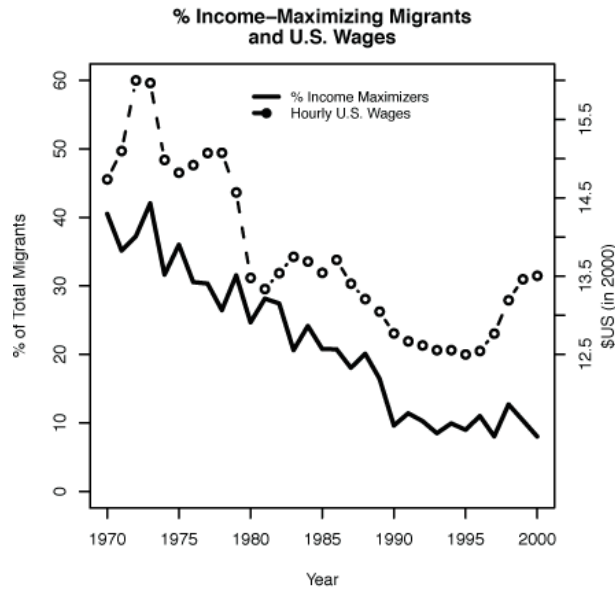
44

Summary of the Results

Income maximizers…

represent the neoclassical economics theory, and

are most prominent in the 1970s…

…when the U.S. wages are at their highest.

Summary of the Results

Income maximizers…

represent the neoclassical economics theory, and

are most prominent in the 1970s…

…when the U.S. wages are at their highest.

Risk diversifiers…

personify the new economics theory, and

gain majority in the 1980s…

…when the Mexican inflation is at its peak.

Summary of the Results

Network migrants…

    symbolize the cumulative causation theory, and

    obtain their highest proportion in 1990s…

       …when U.S. visa availability is at its highest.

Summary of the Results

Network migrants…

symbolize the cumulative causation theory, and

obtain their highest proportion in 1990s…

…when U.S. visa availability is at its highest.

Urban migrants…

are not anticipated in any migration theory, but

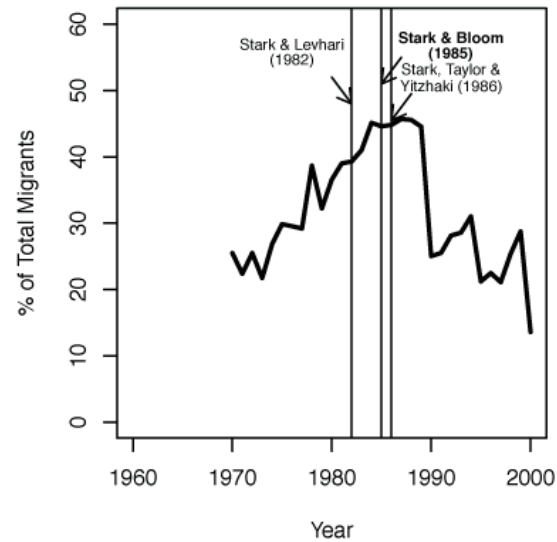loosely connected to globalization arguments; and

become the majority in the 1990s…

…when Mexico-U.S. trade proliferates.
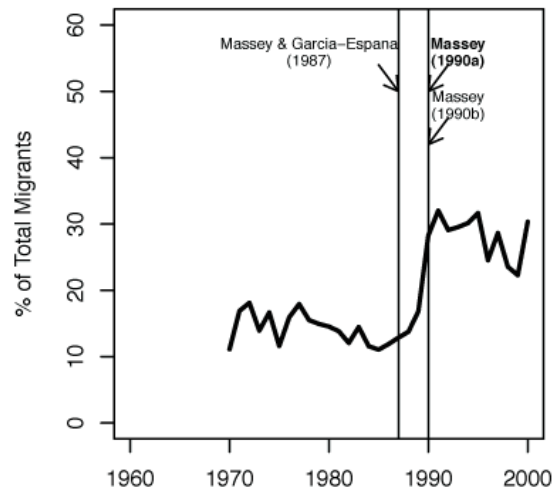
Explaining the Emergence of Theories of Migration?
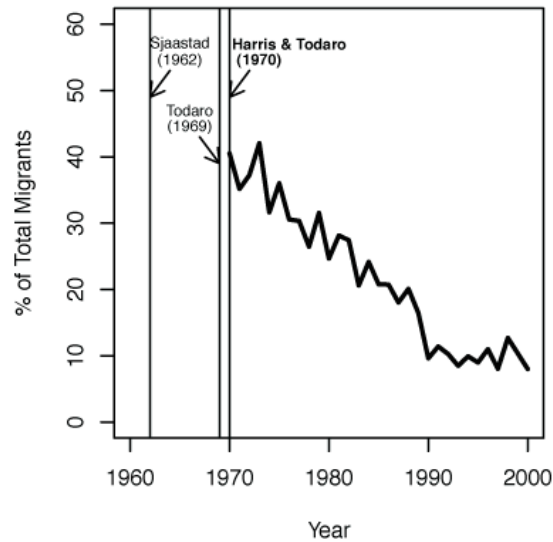
Contributions

This study used cluster analysis to…

    …identify the diverse mechanisms underlying migration, and

    …locate four distinct migrant types in the Mexico-U.S. flows.

Each migrant type…

    …corresponded to a distinct theoretical narrative, and

    …gained prevalence at specific time periods.

The empirical strategy allowed us to…

    …combine various theoretical perspectives, and

    …embrace the diversity of migrants.