

Unity in Diversity?

How Intergroup Contact Can Foster Nation Building

Samuel Bazzi
Boston Univ., NBER

Arya Gaduh
Univ. of Arkansas

Alex Rothenberg
Syracuse Univ.

Maisy Wong
Wharton School

3 September 2020

Mindset Institute

Diversity and Nation Building

- **Uniting Diverse Groups:** a founding principle of many nation-states
- **Nation Building:** Promoting a shared national identity
 - Weaker ethnic attachment, reduced intergroup divisions
 - Building “imagined communities” (Benedict Anderson, 1983)

Diversity and Nation Building

- **Uniting Diverse Groups:** a founding principle of many nation-states
- **Nation Building:** Promoting a shared national identity
 - Weaker ethnic attachment, reduced intergroup divisions
 - Building “imagined communities” (Benedict Anderson, 1983)

Italy has been made; now it remains to make Italians

— Massimo d’Azeglio (1860), quoted in Alesina and Reich (2015)

Diversity and Nation Building

- **Uniting Diverse Groups:** a founding principle of many nation-states
- **Nation Building:** Promoting a shared national identity
 - Weaker ethnic attachment, reduced intergroup divisions
 - Building “imagined communities” (Benedict Anderson, 1983)

Italy has been made; now it remains to make Italians

— Massimo d’Azeglio (1860), quoted in Alesina and Reich (2015)

[T]he most certain prediction that we can make about almost any modern society is that it will be more diverse a generation from now than it is today . . . the central challenge for modern, diversifying societies is to create a new, broader sense of we.

— Putnam (2007)

Intergroup Contact and Nation Building

- **Migration** \implies \uparrow **Local Diversity** $\stackrel{?}{\implies}$ **Nation Building**
 - Negative short-run effects of increases in diversity
(Fearon/Laitin, 2011)
 - Intergroup ties may form + Δ long run preferences through contact
(Allport, 1954)

Intergroup Contact and Nation Building

- **Migration** \implies \uparrow **Local Diversity** $\stackrel{?}{\implies}$ **Nation Building**
 - Negative short-run effects of increases in diversity (Fearon/Laitin, 2011)
 - Intergroup ties may form + Δ long run preferences through contact (Allport, 1954)
- **Difficult to Identify Long Run Impacts of Diversity:**
 - Local diversity often dissipates through sorting, tipping, and segregation (Schelling, 1971)
 - Long-run diversity confounded by geography and endogenous sorting (Michalopoulos, 2012)
 - Fractionalization (F) and polarization (P) may have different effects (Esteban and Ray, 2011)

How Does Intergroup Contact Affect Nation Building?

Evidence from a Large-Scale Policy Experiment

- **Indonesia:** Expansive archipelago, ethnolinguistically diverse
 - Ethnic groups **relatively isolated** from each other historically
 - Regional separatism threatened viability of nation state
 - **Historical Cleavages:** core **Inner Island** vs. periphery **Outer Island**

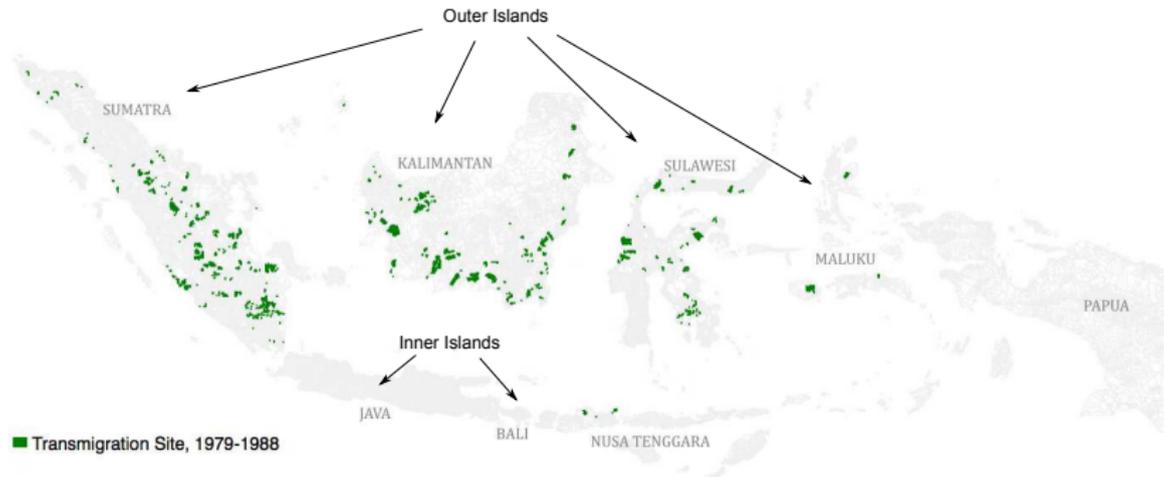


→ **Population Resettlement** as part of nation building policy

A Natural Policy Experiment in Ethnic Mixing

Transmigration: Voluntary Rural-to-Rural Resettlement, 1979–1988

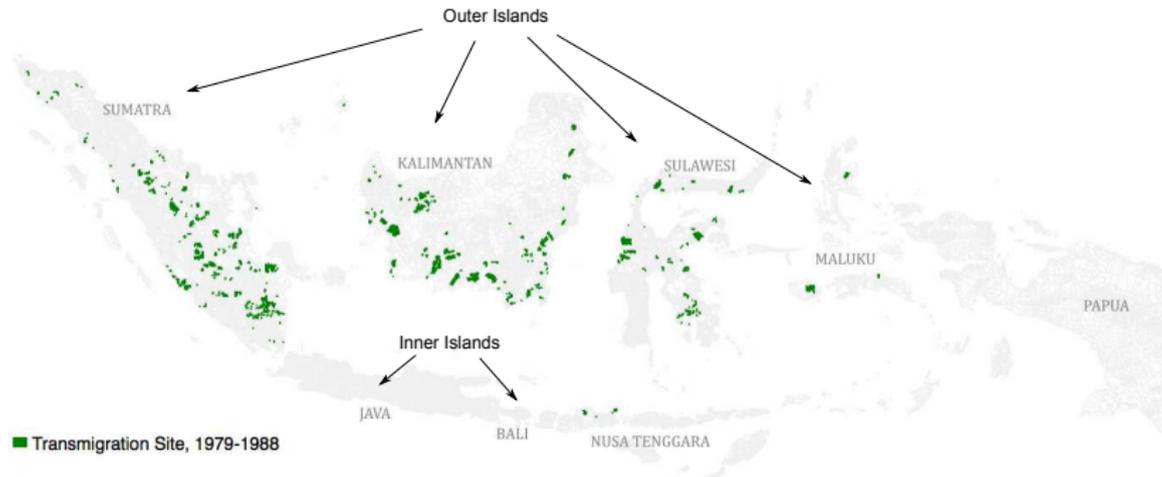
- 2 million migrants from Java/Bali placed in **>900 new villages**
- Each community contained a **mix of Inner and Outer Islanders**
- **Goals:** population redistribution + integration, but very controversial



A Natural Policy Experiment in Ethnic Mixing

Transmigration: Voluntary Rural-to-Rural Resettlement, 1979–1988

- 2 million migrants from Java/Bali placed in **>900 new villages**
- Each community contained a **mix of Inner and Outer Islanders**
- **Goals:** population redistribution + integration, but very controversial

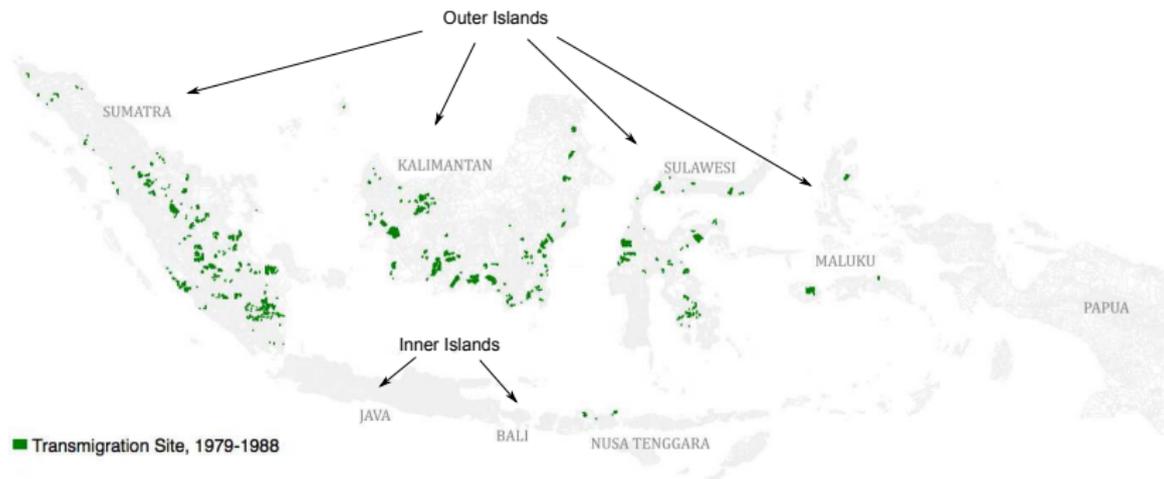


- Conditional quasi-random assignment + migration frictions

A Natural Policy Experiment in Ethnic Mixing

Transmigration: Voluntary Rural-to-Rural Resettlement, 1979–1988

- 2 million migrants from Java/Bali placed in **>900 new villages**
- Each community contained a **mix of Inner and Outer Islanders**
- **Goals:** population redistribution + integration, but very controversial



- Conditional quasi-random assignment + migration frictions
⇒ **persistent, plausibly exogenous variation in LR local diversity**

Measuring Diversity: Fractionalization and Polarization

Fractionalization (F): $F = 1 - \sum_j s_j^2$

- Probability two randomly-drawn individuals come from different (ethno-religious) group
- $F = 1 -$ Herfindahl concentration index
- Higher F \implies Many small groups

Polarization (P): $P = 4 \sum_j s_j^2 (1 - s_j)$

- How group memberships are “clustered” (among a small number of groups > 1)
- Higher P \implies Few larger groups (of similar size)

With more than 2 groups, have **similar F(P)** but **higher/lower P(F)**

Preview of Results

Outcomes ~15–30 Years Later

- **Polarization** increases ethnic attachment
- **Fractionalization** increases national identity and integration

Preview of Results

Outcomes ~15–30 Years Later

- **Polarization** increases ethnic attachment
- **Fractionalization** increases national identity and integration
 - National language, *Bahasa Indonesia*, vs. ethnic language use at home

Preview of Results

Outcomes ~15–30 Years Later

- **Polarization** increases ethnic attachment
- **Fractionalization** increases national identity and integration
 - National language, *Bahasa Indonesia*, vs. ethnic language use at home
 - Intergroup preferences, trust, and public goods participation
 - Identity content of children's names
 - Interethnic marriage

Preview of Results

Outcomes ~15–30 Years Later

- **Polarization** increases ethnic attachment
- **Fractionalization** increases national identity and integration
 - National language, *Bahasa Indonesia*, vs. ethnic language use at home
 - Intergroup preferences, trust, and public goods participation
 - Identity content of children's names
 - Interethnic marriage
 - Village-level public goods, local development, and ethnic conflict

Preview of Results

Outcomes ~15–30 Years Later

- **Polarization** increases ethnic attachment
- **Fractionalization** increases national identity and integration
 - National language, *Bahasa Indonesia*, vs. ethnic language use at home
 - Intergroup preferences, trust, and public goods participation
 - Identity content of children's names
 - Interethnic marriage
 - Village-level public goods, local development, and ethnic conflict
- **Key Mechanisms:** Intergroup Distances
 1. Contact with local neighbors (**residential segregation**)
 2. Interethnic inequality in skills (**economic interaction**)
 3. Predetermined linguistic distance (**cultural differences**)

Preview of Results

Outcomes ~15–30 Years Later

- **Polarization** increases ethnic attachment
- **Fractionalization** increases national identity and integration
 - National language, *Bahasa Indonesia*, vs. ethnic language use at home
 - Intergroup preferences, trust, and public goods participation
 - Identity content of children's names
 - Interethnic marriage
 - Village-level public goods, local development, and ethnic conflict
- **Key Mechanisms:** Intergroup Distances
 1. Contact with local neighbors (**residential segregation**)
 2. Interethnic inequality in skills (**economic interaction**)
 3. Predetermined linguistic distance (**cultural differences**)
- **Model of Identity Choices:** Evolutionary Game Theory
 - ⇒ **Growth of national identity** via contact amidst diversity

Contributions and Related Literature

1. Nation Building (Policy) amidst Diversity and Migration

(Alesina/Reich, 2015; Bandiera et al, 2016; Blouin/Mukand, 2016; Clots-Figuera/Masella, 2013; Fouka, 2016; Laitin/Ramachandran, 2016; Miguel, 2004; Okunogbe, 2015)

2. Contact Hypothesis

(Green et al, 2018; Lowe, 2018; Rao, 2018; ...)

3. Intergenerational Process of Cultural Change in Diverse Societies

(Algan et al, 2016; Bisin et al; Clingingsmith et al, 2009; Desmet et al, 2017; Fernandez, 2011)

Key Contributions to this Growing Literature:

- Policy-induced, long-run variation in diversity with limited sorting
- Long-run changes in revealed preference for integration
- New shared identity, distinct from minority assimilation or conformity
- Clarifying distinct effects of polarization and fractionalization
- Conditions that facilitate integration amidst diversity

Roadmap

Introduction

Background: Diversity and the Transmigration Program

Data: Diversity and Nation Building Outcomes

Results: Diversity, Socialization, and Identity

- National Language Use at Home

- Mechanisms

- Other Outcomes

A Model of Growth in National Identity

Conclusion

Roadmap

Introduction

Background: Diversity and the Transmigration Program

Data: Diversity and Nation Building Outcomes

Results: Diversity, Socialization, and Identity

- National Language Use at Home

- Mechanisms

- Other Outcomes

A Model of Growth in National Identity

Conclusion

Diversity and the Problem of Unity

- Indonesia: **ethnolinguistically diverse** island nation
 - > 1,000 ethnicities, 700 languages **but** living in **homogenous villages**
(median *village* $F = 0.05$, *national* $F = 0.7$)
 - 14 native Inner-Island groups: Java, Sunda, Bali, Madura largest
 - 900+ native Outer-Island groups: several large ones on each island
(biggest groups: Minang, Bugis, Aceh, Batak, Banjar, Dayak, Toraja)
 - typical Inner–Outer linguistic differences \approx German vs. French 

Diversity and the Problem of Unity

- Indonesia: **ethnolinguistically diverse** island nation
 - > 1,000 ethnicities, 700 languages **but** living in **homogenous villages**
(median *village F* = 0.05, *national F* = 0.7)
 - 14 native Inner-Island groups: Java, Sunda, Bali, Madura largest
 - 900+ native Outer-Island groups: several large ones on each island
(biggest groups: Minang, Bugis, Aceh, Batak, Banjar, Dayak, Toraja)
 - typical Inner–Outer linguistic differences \approx German vs. French 
- **Nation building** was an important **concern for policymakers**
 - 1928 Youth Pledge: a pre-independence declaration of Indonesian unity
 - “Unity in diversity”: national motto (also in E.U.), enshrined in coat of arms



Diversity and the Problem of Unity

- Indonesia: **ethnolinguistically diverse** island nation
 - > 1,000 ethnicities, 700 languages **but** living in **homogenous villages**
(median *village F* = 0.05, *national F* = 0.7)
 - 14 native Inner-Island groups: Java, Sunda, Bali, Madura largest
 - 900+ native Outer-Island groups: several large ones on each island
(biggest groups: Minang, Bugis, Aceh, Batak, Banjar, Dayak, Toraja)
 - typical Inner–Outer linguistic differences \approx German vs. French 
- **Nation building** was an important **concern for policymakers**
 - 1928 Youth Pledge: a pre-independence declaration of Indonesian unity
 - “Unity in diversity”: national motto (also in E.U.), enshrined in coat of arms
- Adoption of **the national language**: a key marker of identity



Identity Building Through Language Policy

[T]he more [people] learned to express themselves in Indonesian, the more conscious they became of the ties which linked them.

— Alisjahbana, 1962

- ***Bahasa Indonesia***: a key choice in the 1928 Declaration of Unity
 - Historical lingua franca: *Malay*
 - Not language of the plurality (Javanese \approx 40% of pop.)
- **Rapid growth in national language use**
 - 1930s: roughly 5% of the population able to speak Indonesian
 - Today: \approx 93% can speak
 - 18% main language at home
 - 95% main language at home in Jakarta
- *Asia Barometer*: People who mostly speak *Bahasa Indonesia* at home
 \implies 15% \uparrow attachment to Indonesian rather than own ethnic identity

Identity Building Through Language Policy



Pew Research Center *Global Attitudes & Trends*

HOME

U.S. POLITICS

MEDIA & NEWS

SOCIAL TRENDS

RELIGION

INTERNET & TECH

PUBLICATIONS

TOPICS

DATASETS

QUESTION SEARCH

GLOBAL INDICATOR

FEBRUARY 1, 2017

What It Takes to Truly Be ‘One of Us’

In U.S., Canada, Europe, Australia and Japan, publics say language matters more to national identity than birthplace

BY BRUCE STOKES

The tide of **people moving across the world**,
be they immigrants or refugees, has sparked

Relatively few say national identity is

Majorities in every country surveyed say it is very important to speak the dominant language to be considered truly a national of that land. This includes a median of 77% in Europe and majorities in Japan (70%), the U.S. (70%), Australia (69%) and Canada (59%)

— Source: Pew Research Center (Feb 2017)

Resettlement as Part of Nation Building under Suharto

- **Transmigration:** large-scale resettlement in late 1970s
 - Concerns about Density: Java/Bali 66% of pop., 7% of land
 - Goals: population redistribution, food security, nation building
 - Budget: \$6.6 billion USD, funded by oil revenue windfall

By way of transmigration, we will try to . . . integrate all the ethnic groups into one nation, the Indonesian nation. The different ethnic groups will in the long run disappear because of integration and there will be one kind of man, Indonesian.

— Martono, Minister of Transmigration, 1985

Resettlement as Part of Nation Building under Suharto

- **Transmigration:** large-scale resettlement in late 1970s
 - Concerns about Density: Java/Bali 66% of pop., 7% of land
 - Goals: population redistribution, food security, nation building
 - Budget: \$6.6 billion USD, funded by oil revenue windfall

By way of transmigration, we will try to . . . integrate all the ethnic groups into one nation, the Indonesian nation. The different ethnic groups will in the long run disappear because of integration and there will be one kind of man, Indonesian.

— Martono, Minister of Transmigration, 1985

- Skeptics viewed program as vehicle for “Javanization” of Outer Islands (Charras et al, 1993; Levang, 1995; Schiller & Ganang, 2002)
- Popular fears of violent conflict between Inner and Outer Islanders (lots of anecdotes + claims in Fearon & Laitin, 2011 re Papua)

Transmigration Program Implementation

- **Selecting Sites:**

- New Villages and Farms: created on previously uncleared federal land
- Site Selection: based on geographic and agroclimatic features (x)
(topography, soil quality, water access, weather, transport access)

Transmigration Program Implementation

- **Selecting Sites:**

- New Villages and Farms: created on previously uncleared federal land
- Site Selection: based on geographic and agroclimatic features (x) (topography, soil quality, water access, weather, transport access)

- **Designing New Settlements**

- Carrying Capacity: based on land quality and quantity
- Slots for Local Outer-Island natives (APPDT): *de jure*, 10–30%
de facto, some settlements included as high as 50–80%
- Lottery: allocates house + 2 ha farm plots, ownership after 5-10 yrs
- Public Institutions (schools, gov't offices) identical in all settlements

Transmigration Program Implementation

- **Selecting Sites:**

- New Villages and Farms: created on previously uncleared federal land
- Site Selection: based on geographic and agroclimatic features (x) (topography, soil quality, water access, weather, transport access)

- **Designing New Settlements**

- Carrying Capacity: based on land quality and quantity
- Slots for Local Outer-Island natives (APPDT): *de jure*, 10–30%
de facto, some settlements included as high as 50–80%
- Lottery: allocates house + 2 ha farm plots, ownership after 5-10 yrs
- Public Institutions (schools, gov't offices) identical in all settlements

- **Transmigrant Households**

- Voluntary Participation: married, farmers, household head aged 20-40 (prior schooling and ethnic names similar to stayers in rural Java/Bali)

Advertising the Transmigration Program



A bright and vigorous future, together we move towards a joyous life.

Lack of a Systematic Assignment Mechanism

- Transmigrants sent from 4 transit camps (x), could not choose destinations
 - Little known about destinations pre-departure; 85% did not know local ethnic group (Kebeschull, 1986 camp survey)



- plan-as-you-proceed: “we would just ship out groups of transmigrants as they showed up in transit camps”

Lack of a Systematic Assignment Mechanism

- Transmigrants sent from 4 transit camps (x), could not choose destinations
 - Little known about destinations pre-departure; 85% did not know local ethnic group (Kebeschull, 1986 camp survey)



- plan-as-you-proceed: “we would just ship out groups of transmigrants as they showed up in transit camps”
- **Two Sources of Identifying Variation:**
 1. Inner-Island Ethnic Diversity: Coincidental timing of arrival in transit camps, ethnic mixing in camps, and opening of villages

Lack of a Systematic Assignment Mechanism

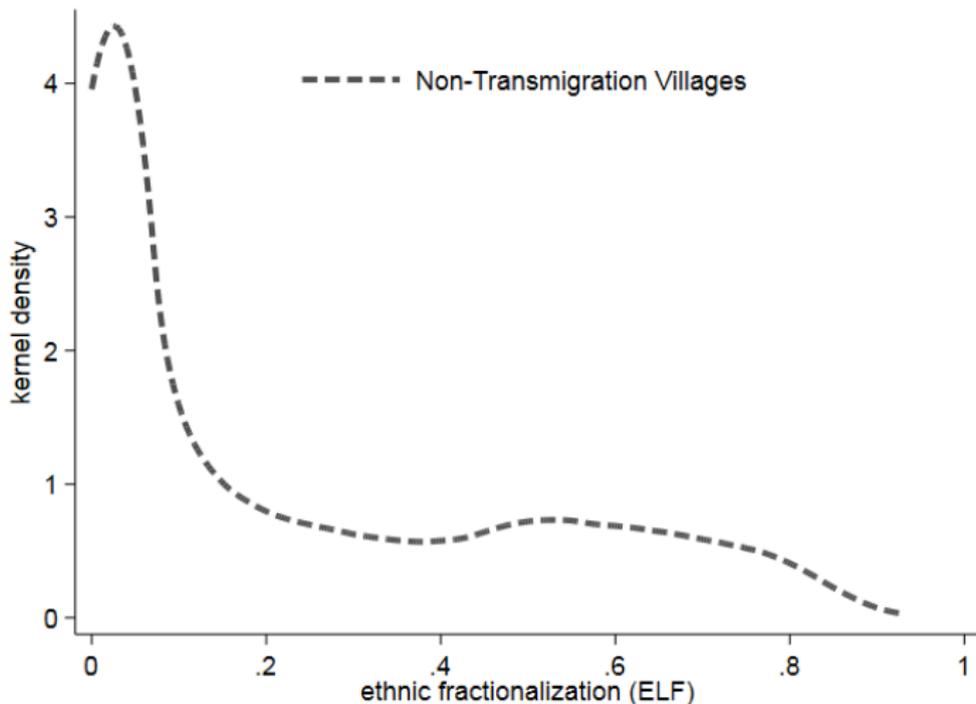
- Transmigrants sent from 4 transit camps (x), could not choose destinations
 - Little known about destinations pre-departure; 85% did not know local ethnic group (Kebeschull, 1986 camp survey)



- plan-as-you-proceed: “we would just ship out groups of transmigrants as they showed up in transit camps”
- **Two Sources of Identifying Variation:**
 1. Inner-Island Ethnic Diversity: Coincidental timing of arrival in transit camps, ethnic mixing in camps, and opening of villages
 2. Outer-Island Ethnic Diversity: Variation by year in APPDT quotas

What Does the Program Buy Us?

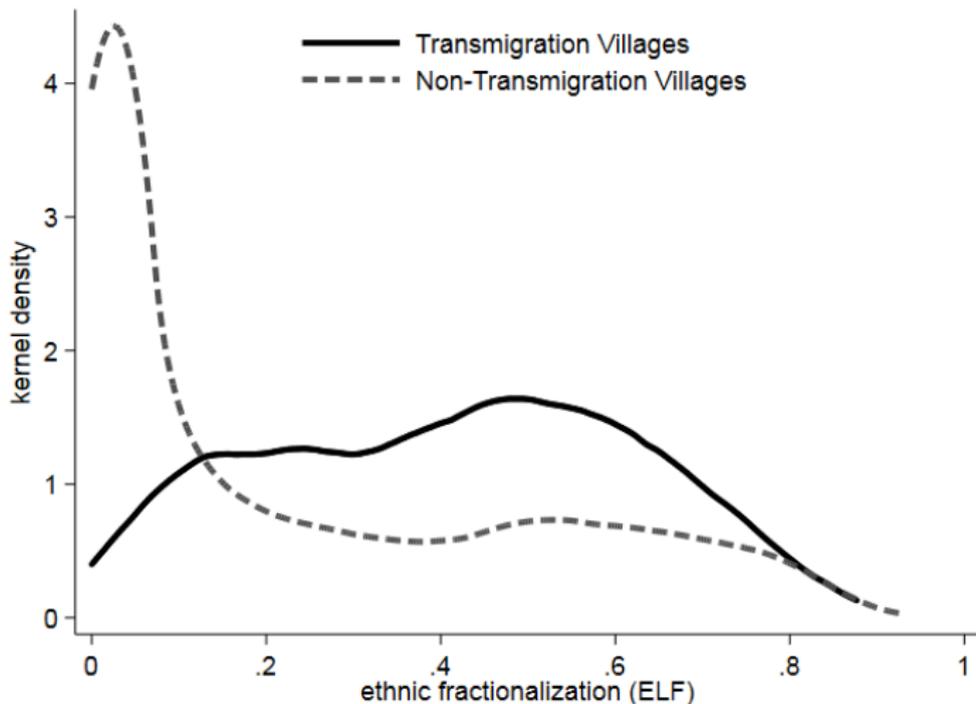
Persistent Continuum of Village-Level Diversity



fractionalization: $F = 1 - \sum_j s_j^2$

What Does the Program Buy Us?

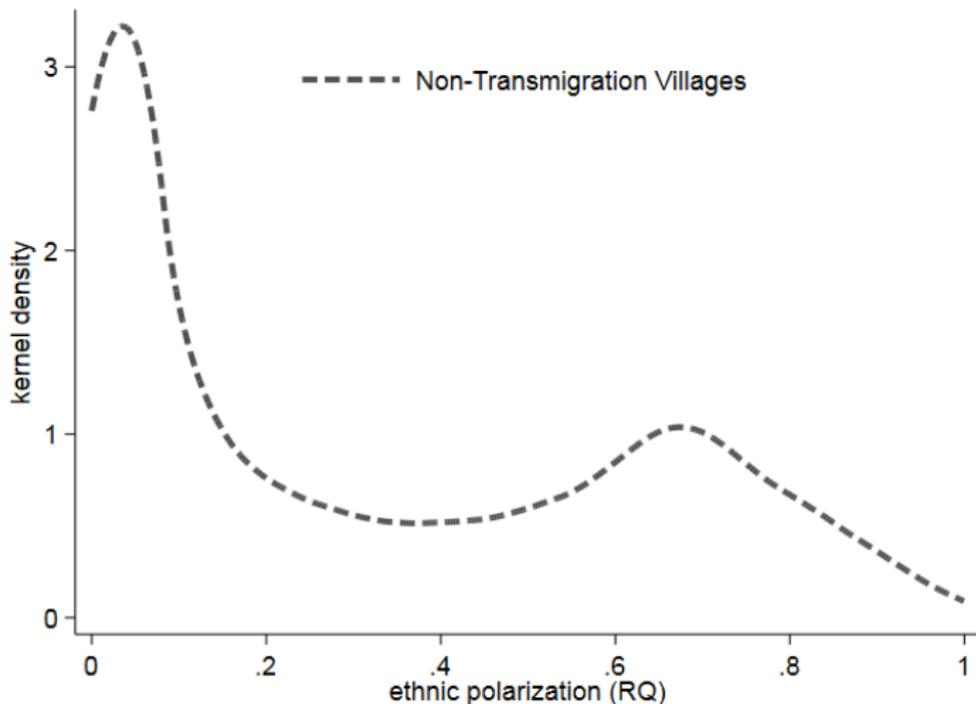
Persistent Continuum of Village-Level Diversity



fractionalization: $F = 1 - \sum_j s_j^2$

What Does the Program Buy Us?

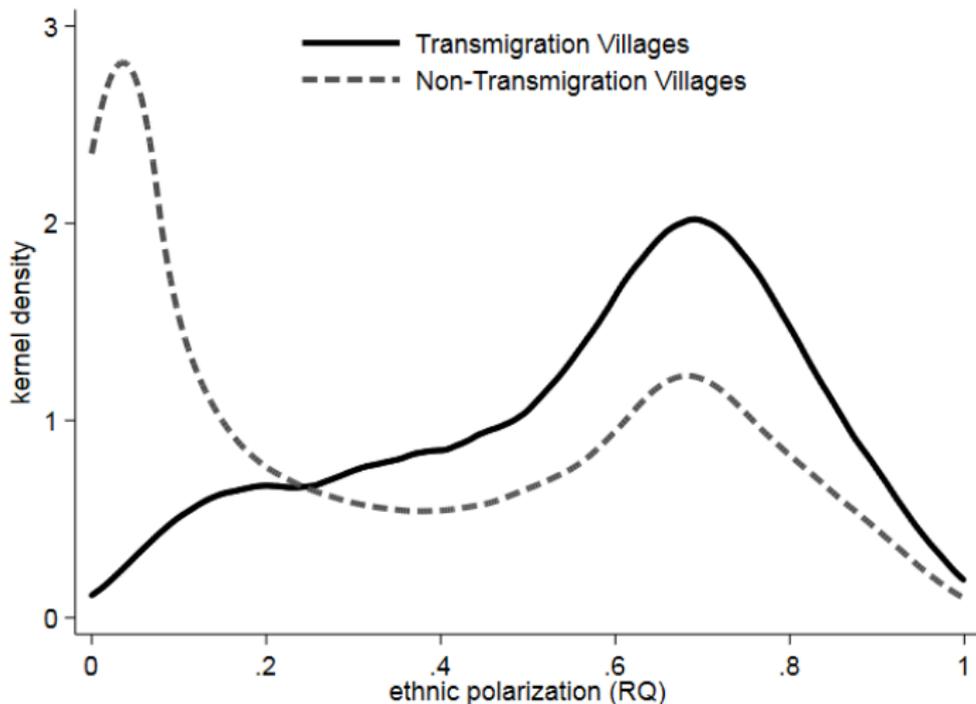
Persistent Continuum of Village-Level Diversity



polarization: $P = 4 \sum_j s_j^2 (1 - s_j)$

What Does the Program Buy Us?

Persistent Continuum of Village-Level Diversity

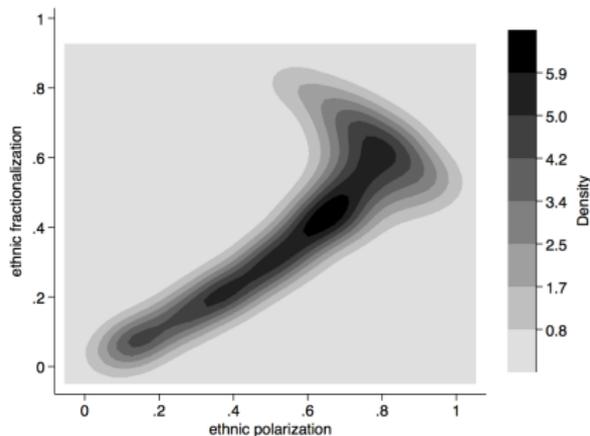


polarization: $P = 4 \sum_j s_j^2 (1 - s_j)$

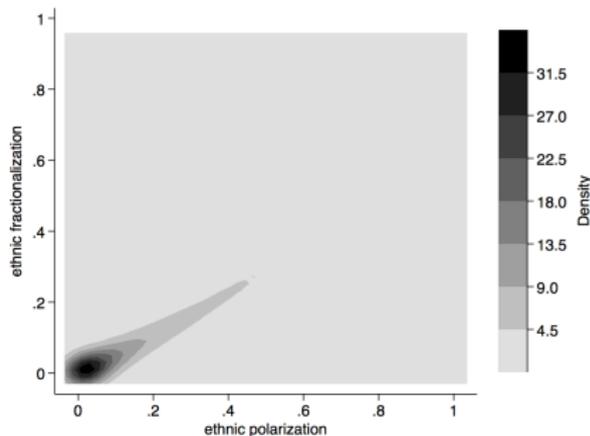
What Does the Program Buy Us?

Persistent Continuum of Village-Level Diversity

(a) Transmigration Villages

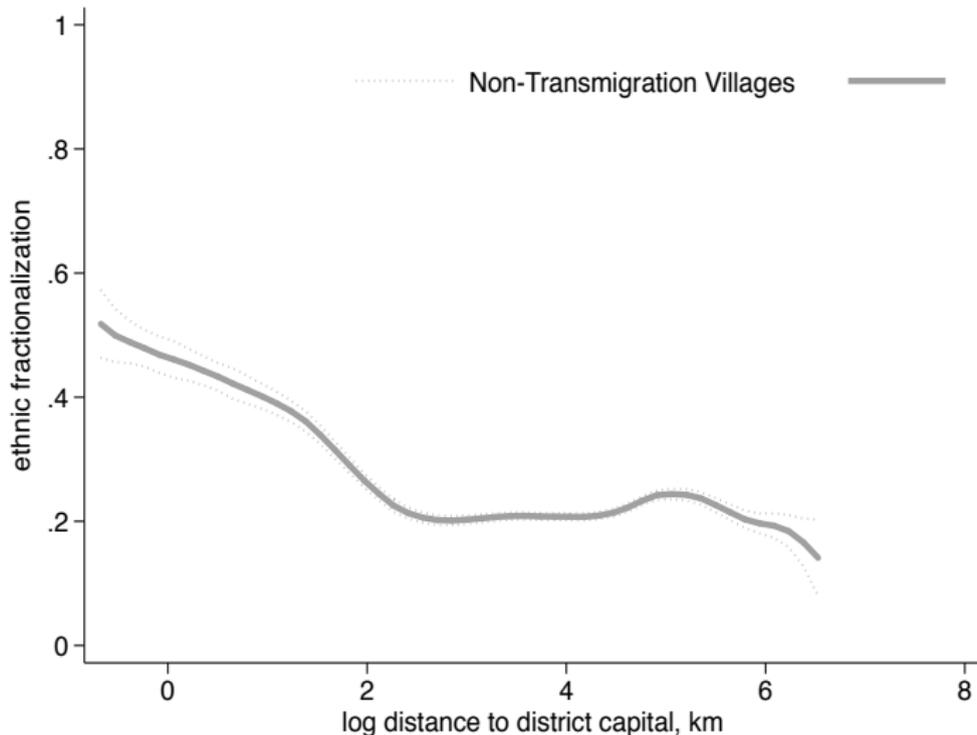


(b) Non-Transmigration Villages



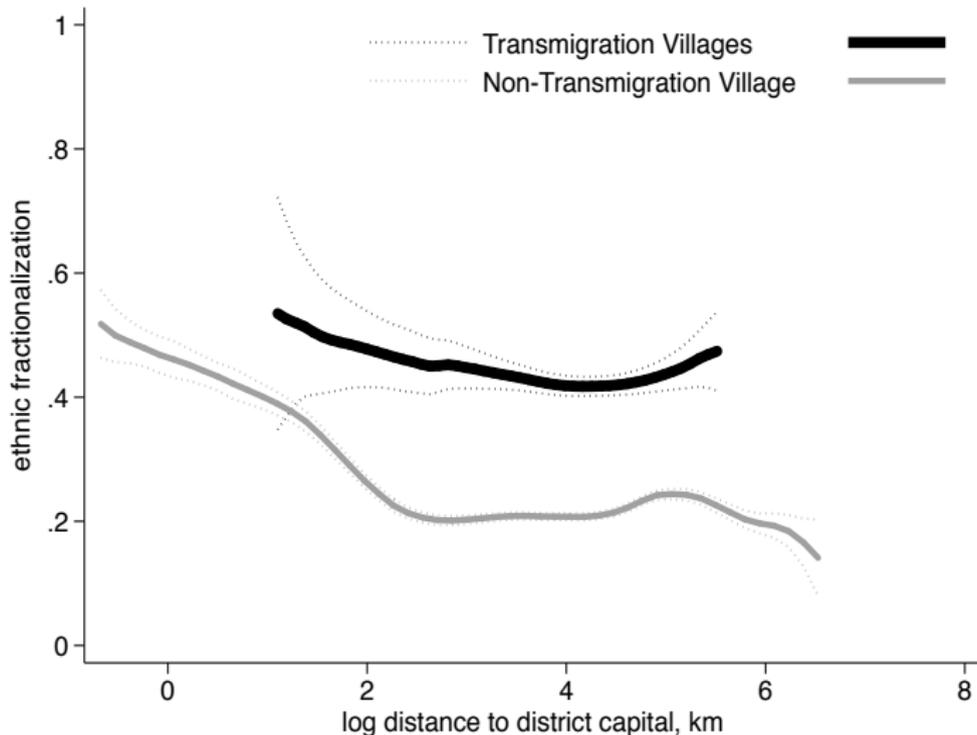
What Does the Program Buy Us?

Diversity More Exogenous w.r.t. Natural Advantages



What Does the Program Buy Us?

Diversity More Exogenous w.r.t. Natural Advantages



Roadmap

Introduction

Background: Diversity and the Transmigration Program

Data: Diversity and Nation Building Outcomes

Results: Diversity, Socialization, and Identity

- National Language Use at Home

- Mechanisms

- Other Outcomes

A Model of Growth in National Identity

Conclusion

Diversity, Names, Language, and Marriage

- **2010 Population Census:** Universal coverage
 - Ethnicity: Each individual's self reported choice
 - Language Use at Home in 2010
 - Full Name in 2010
 - Interethnic Marriage status of household head
 - District of birth

Diversity, Names, Language, and Marriage

- **2010 Population Census:** Universal coverage
 - Ethnicity: Each individual's self reported choice
 - Language Use at Home in 2010
 - Full Name in 2010
 - Interethnic Marriage status of household head
 - District of birth

- **Native Languages:** map Census ethnicities to *Ethnologue*

Diversity, Names, Language, and Marriage

- **2010 Population Census:** Universal coverage
 - Ethnicity: Each individual's self reported choice
 - Language Use at Home in 2010
 - Full Name in 2010
 - Interethnic Marriage status of household head
 - District of birth
 - **Native Languages:** map Census ethnicities to *Ethnologue*
 - **Diversity:** F , P
 - Shapley decomp.: ~50% of variation from inner-outer ethnic divisions
~50% of variation from inner-inner ethnic divisions
- ⇒ Diversity due to program: (1) APPDT slots, (2) mixing in transit camps

Diversity, Names, Language, and Marriage

- **2010 Population Census:** Universal coverage
 - Ethnicity: Each individual's self reported choice
 - Language Use at Home in 2010
 - Full Name in 2010
 - Interethnic Marriage status of household head
 - District of birth
- **Native Languages:** map Census ethnicities to *Ethnologue*
- **Diversity:** F , P
 - Shapley decomp.: ~50% of variation from inner-outer ethnic divisions
~50% of variation from inner-inner ethnic divisions

⇒ Diversity due to program: (1) APPDT slots, (2) mixing in transit camps
- **Ethnic Residential Segregation**
own village → neighborhood → next-door neighbors

National Language as Vehicle for Nation Building

Indonesian is the most clearly defined and regularly experienced aspect of Indonesian national culture. . .

— Simpson, 2007 “Language and National Identity in Asia”

Indonesian has also become positively valued as the primary shared component of the country's emerging national identity.

— Heryanto, 1995

National Language as Vehicle for Nation Building

Using the *Indonesia Family Life Survey*, we relate individual outcomes in 2014 to their parents' choices from their former household in 1997:

$$y_{ij}^{14} = \alpha + \eta \mathbb{1} \left\{ \text{Indonesian at home}_{ij}^{97} \right\} + \mathbf{x}'_{ij} \boldsymbol{\delta} + \theta_j + \varepsilon_{ij}$$

	Dependent Variable as Adult in 2014:			
	Speaks Indonesian at Home	Changes Ethnicity from 1997	In Interethnic Marriage	Trust Other Ethnic Groups (z-score)
Panel A: Baseline	(1)	(2)	(3)	(4)
Indonesian was Primary Language at Home as Child in 1997	0.156 (0.022)***	0.062 (0.019)***	0.053 (0.023)**	0.148 (0.054)***
Dependent Variable Mean	0.369	0.114	0.103	0.00
Age, Gender, Education Fixed Effects	Yes	Yes	Yes	Yes
Village Fixed Effects	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by villages in parentheses. */**/** denotes significant at the 10/5/1 percent significance.

Childhood Indonesian use ~ weaker ethnic attachment

National Language as Vehicle for Nation Building

	Dependent Variable as Adult in 2014:			
	Speaks Indonesian at Home	Changes Ethnicity from 1997	In Interethnic Marriage	Trust Other Ethnic Groups (z-score)
Panel B: Adding Parental Inter-marriage	(1)	(2)	(3)	(4)
Indonesian was Primary Language at Home as Child in 1997	0.151 (0.022)***	0.045 (0.019)**	0.046 (0.023)**	0.131 (0.054)**
Parents from Different Ethnic Groups	0.053 (0.021)**	0.177 (0.030)***	0.092 (0.031)***	0.160 (0.055)***
Number of Individuals	8,623	6,594	5,628	8,236
Dependent Variable Mean	0.369	0.114	0.103	0.00
Age, Gender, Education Fixed Effects	Yes	Yes	Yes	Yes
Village Fixed Effects	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by villages in parentheses. */**/** denotes significant at the 10/5/1 percent significance.

Effects not driven entirely by children of interethnic marriages

Roadmap

Introduction

Background: Diversity and the Transmigration Program

Data: Diversity and Nation Building Outcomes

Results: Diversity, Socialization, and Identity

- National Language Use at Home

- Mechanisms

- Other Outcomes

A Model of Growth in National Identity

Conclusion

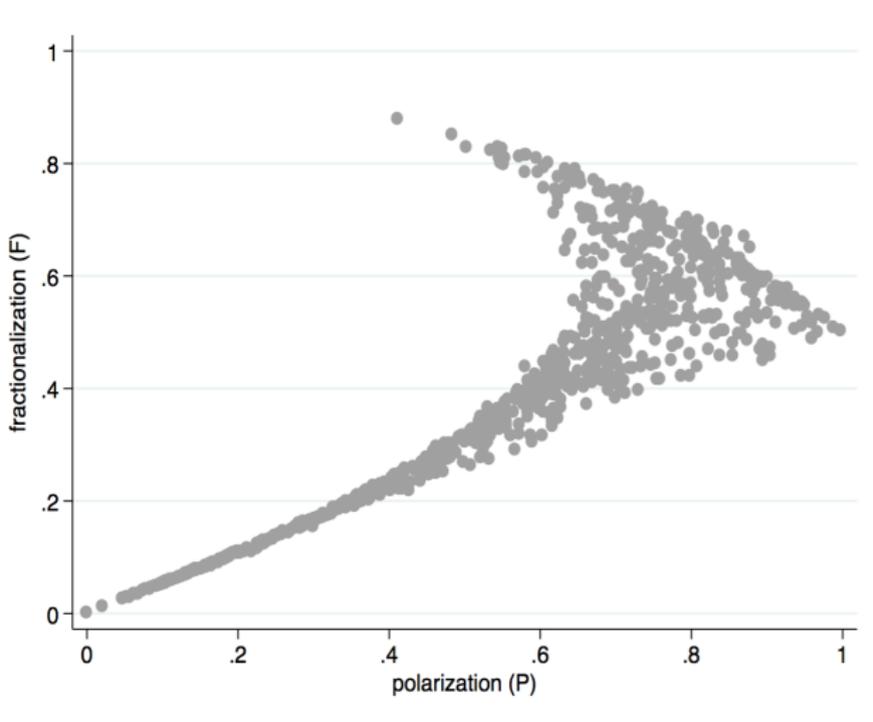
Empirical Strategy

$$\begin{aligned}y_{iv} &= \alpha + g(\text{diversity}_v) + \mathbf{x}'_{iv}\boldsymbol{\beta} + \varepsilon_{iv} \\ &= \alpha + \beta_F F_v + \beta_P P_v + \mathbf{x}'_{iv}\boldsymbol{\beta} + \varepsilon_{iv}\end{aligned}$$

- y_{iv} : nation building outcome for individual i in village v
- \mathbf{x} : natural advantages used to select sites + fixed effects
→ individuals of same ethnicity e , same origin, same age, ...
- F_v : fractionalization
 - $\beta_F < 0$: reduced cooperation, cultural entrenchment
 - $\beta_F > 0$: cultural learning, reduced prejudice, greater integration
- P_v : polarization
 - $\beta_P < 0$: conflict, stronger ethnic attachment, weaker integration
 - $\beta_P > 0$: group cohesion, stronger enforcement, greater integration

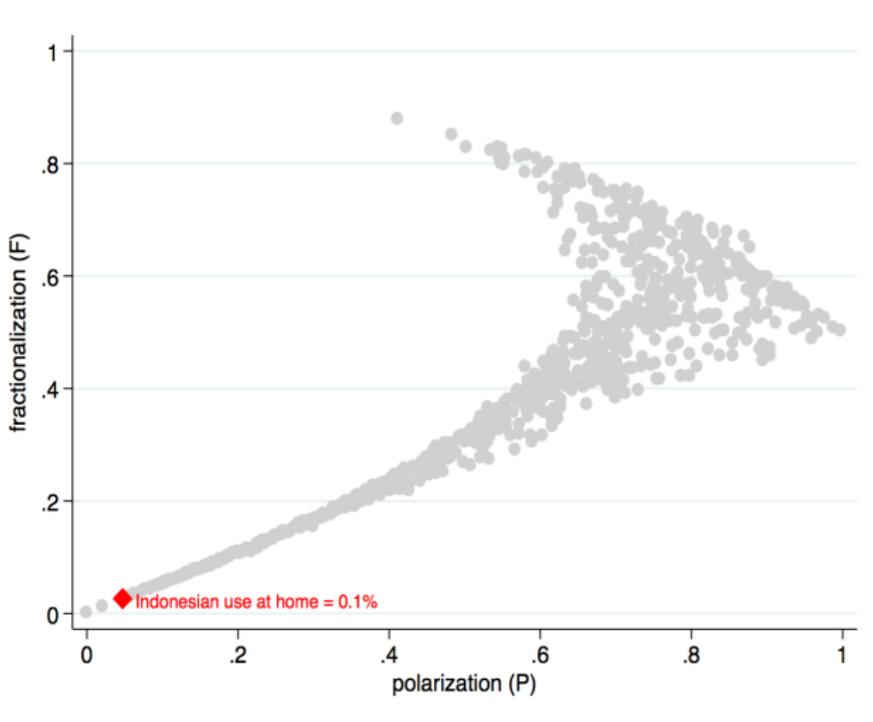
Diversity and National Language Use At Home

Motivating Examples: Increasing F and P



Diversity and National Language Use At Home

Motivating Examples: Increasing F and P

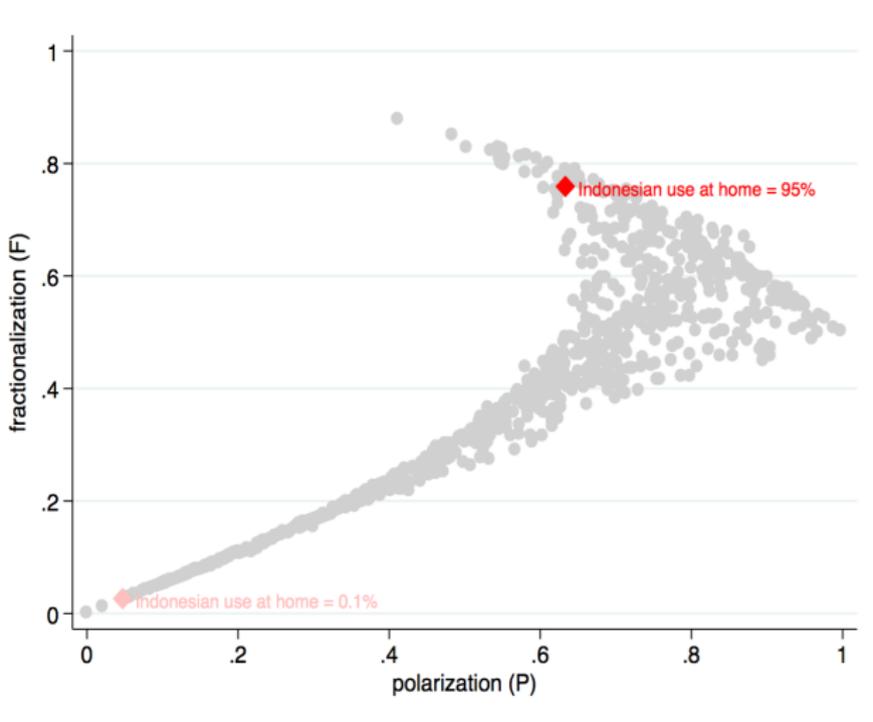


Tanjung Damai village in Riau Province

- 98.8% Javanese, 1.2% Melayu Riau
- $F = 0.02$, $P = 0.05$

Diversity and National Language Use At Home

Motivating Examples: Increasing F and P

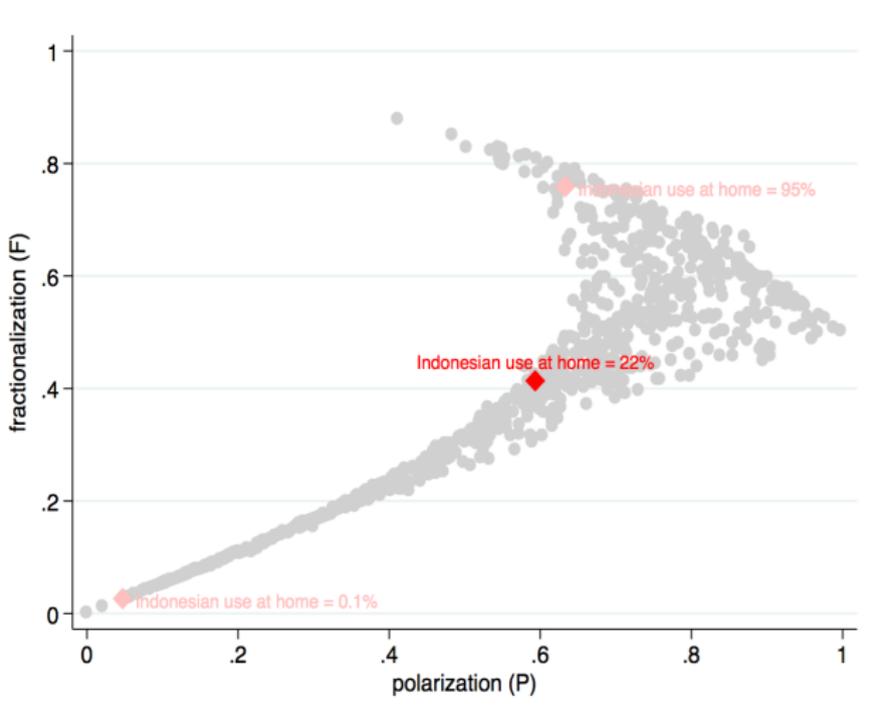


Tanjung Gading village in Lampung Province

- 43 ethnic groups; 42% Javanese, 21% Banten, 11% Lampung, ...
- $F = 0.76$, $P = 0.63$

Diversity and National Language Use At Home

Motivating Examples: Increasing F and P

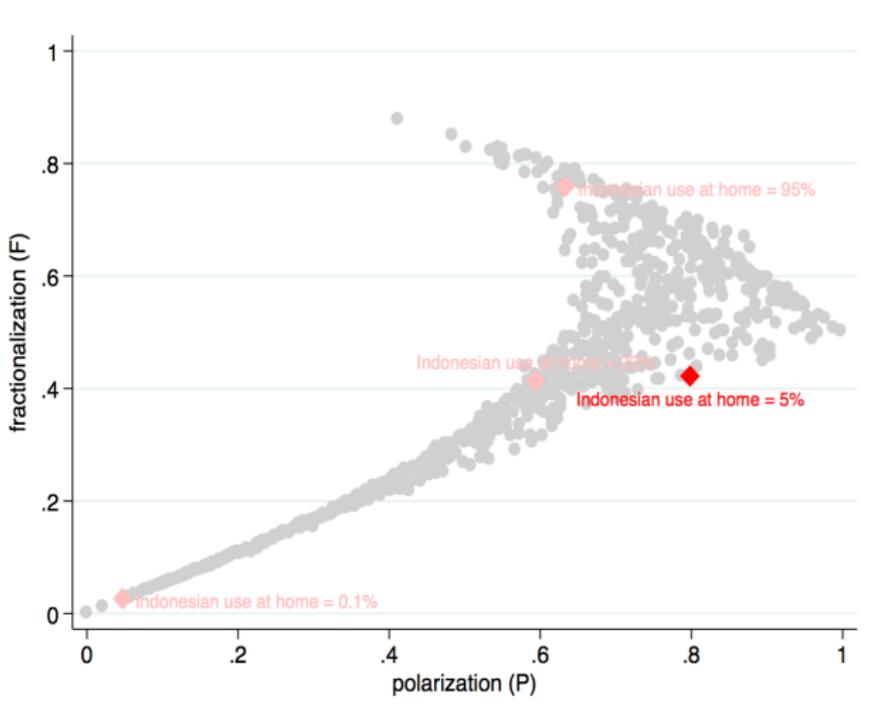


Bukit Kemuning village in Province

- 76% Javanese, 7% Minangkabau, 7% Batak Toba, ...
- $F = 0.41$, $P = 0.59$

Diversity and National Language Use At Home

Motivating Examples: Increasing F and P

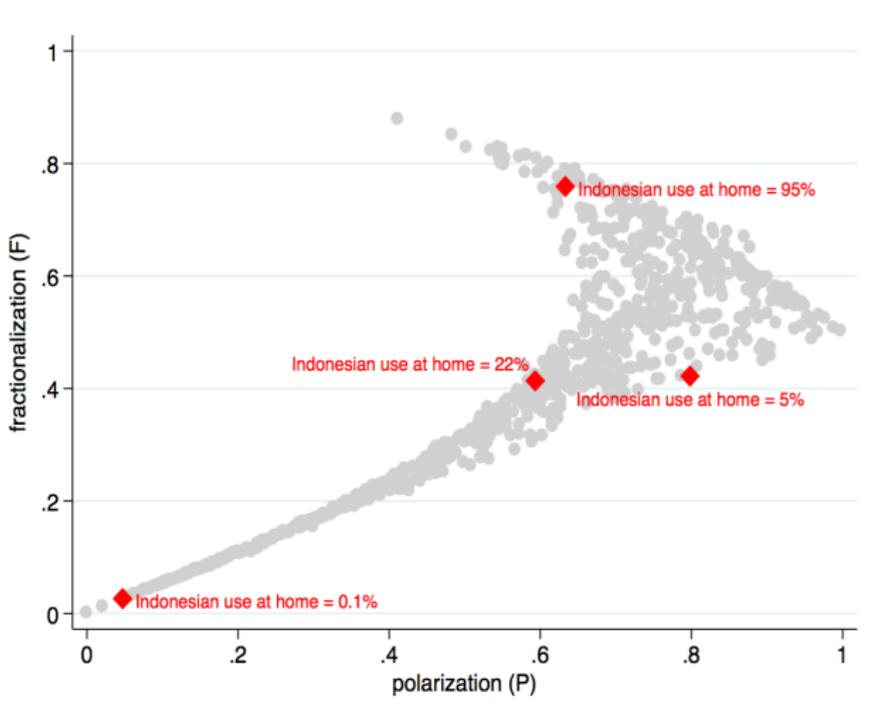


Wonodadi village in South Sumatra Province

- 71.2% Javanese, 27% Sunda, 1.5% Melayu
- $F = 0.42$, $P = 0.80$

Diversity and National Language Use At Home

Motivating Examples: Increasing F and P



Threats to Identification and IV Strategy

$$y_{iv} = \alpha + \beta_F F_v + \beta_P P_v + \mathbf{x}'_{iv} \boldsymbol{\beta} + \varepsilon_{iv}$$

- $(F, P) = \text{ex ante assignment} + \text{ex post sorting}$

Threats to Identification and IV Strategy

$$y_{iv} = \alpha + \beta_F F_v + \beta_P P_v + \mathbf{x}'_{iv} \boldsymbol{\beta} + \varepsilon_{iv}$$

- $(F, P) = \text{ex ante assignment} + \text{ex post sorting}$
- Direct tests to address endogenous assignment and sorting

Threats to Identification and IV Strategy

$$y_{iv} = \alpha + \beta_F F_v + \beta_P P_v + \mathbf{x}'_{iv} \boldsymbol{\beta} + \varepsilon_{iv}$$

- $(F, P) = \text{ex ante assignment} + \text{ex post sorting}$
- Direct tests to address endogenous assignment and sorting
- We also propose **program-based instruments** | carrying capacity \mathbf{x}_v
 1. **number of Inner-Island transmigrants assigned**
 \implies Inner-Island ethnic share (**inner-outer division**)
 2. **ethnic group shares among those born in Java/Bali**
 \implies F and P within Inner Islanders (**inner-inner divisions**)
- **Intuition:** isolate portion of (F_v, P_v) driven by *ex ante* assignment

▶ IV relevance

▶ probing validity

Diversity in Transmigration Villages Uncorrelated with Natural Advantages and Pre-1979 Development

	District-Level Population Characteristics, 1978						
	distance to district cap. (1)	distance to major road (2)	total population (3)	Indonesian use at home (4)	television ownership (5)	agriculture empl. share (6)	wage empl. share (7)
Transmigration Villages							
ethnic fractionalization	0.146 (0.528)	0.019 (0.041)	-0.267 (0.351)	0.034 (0.038)	-0.005 (0.022)	0.028 (0.044)	-0.019 (0.027)
ethnic polarization	-0.241 (0.432)	-0.008 (0.031)	-0.178 (0.254)	-0.020 (0.024)	0.008 (0.016)	-0.034 (0.032)	0.047** (0.021)
Number of Villages	817	817	817	817	817	817	817
Dependent Variable Mean	4.122	0.079	12.505	0.072	0.069	0.780	0.121
R ²	0.014	0.011	0.240	0.473	0.087	0.032	0.034

Notes: */**/** denotes significance at the 10/5/1 percent significance levels.

Diversity in Transmigration Villages Uncorrelated with Natural Advantages and Pre-1979 Development

	District-Level Population Characteristics, 1978						
	distance to district cap. (1)	distance to major road (2)	total population (3)	Indonesian use at home (4)	television ownership (5)	agriculture empl. share (6)	wage empl. share (7)
Transmigration Villages							
ethnic fractionalization	0.146 (0.528)	0.019 (0.041)	-0.267 (0.351)	0.034 (0.038)	-0.005 (0.022)	0.028 (0.044)	-0.019 (0.027)
ethnic polarization	-0.241 (0.432)	-0.008 (0.031)	-0.178 (0.254)	-0.020 (0.024)	0.008 (0.016)	-0.034 (0.032)	0.047** (0.021)
Number of Villages	817	817	817	817	817	817	817
Dependent Variable Mean	4.122	0.079	12.505	0.072	0.069	0.780	0.121
R ²	0.014	0.011	0.240	0.473	0.087	0.032	0.034
Non-Transmigration Villages							
ethnic fractionalization	-2.166*** (0.288)	-0.048*** (0.016)	-0.436* (0.233)	0.165*** (0.051)	0.109** (0.043)	-0.166* (0.086)	0.114** (0.047)
ethnic polarization	1.465*** (0.207)	0.027** (0.012)	0.294* (0.163)	-0.043 (0.034)	-0.053* (0.029)	0.109* (0.059)	-0.054* (0.032)
Number of Villages	26,119	29,158	22,400	22,400	22,400	22,400	22,400
Dependent Variable Mean	3.517	0.069	12.667	0.084	0.072	0.759	0.133
R ²	0.067	0.136	0.235	0.329	0.146	0.077	0.069

Notes: */**/** denotes significance at the 10/5/1 percent significance levels.

Diversity and National Language Use At Home

Village and Individual-Level Regressions

	<i>Dep. Var.: National Language Use at Home</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
	Village-Level			Individual-Level		
ethnic fractionalization	0.296*** (0.041)		0.637*** (0.073)	0.671*** (0.075)	0.499*** (0.057)	0.377*** (0.051)
ethnic polarization		0.086*** (0.030)	-0.362*** (0.051)	-0.392*** (0.057)	-0.302*** (0.041)	-0.184*** (0.038)
Number of Villages	817	817	817	817	817	817
Number of Individuals	-	-	-	1,800,499	1,800,499	1,800,499
Dependent Variable Mean	0.144	0.144	0.144	0.154	0.154	0.154
R ²	0.379	0.303	0.437	0.114	0.221	0.280
Island FE, Predetermined Controls (x)	✓	✓	✓	✓	✓	✓
Ethnicity, Age, Relation, Gender FE					✓	✓
Birth District, Current District FE						✓

Notes: */**/** denotes significance at the 10/5/1 percent significance levels. Standard errors clustered by district.

- One SD $\uparrow F \implies$ 12.9 p.p. \uparrow *HomeIndo*
- One SD $\uparrow P \implies$ 8.1 p.p. \downarrow *HomeIndo*

Diversity and National Language Use At Home

IV Regressions

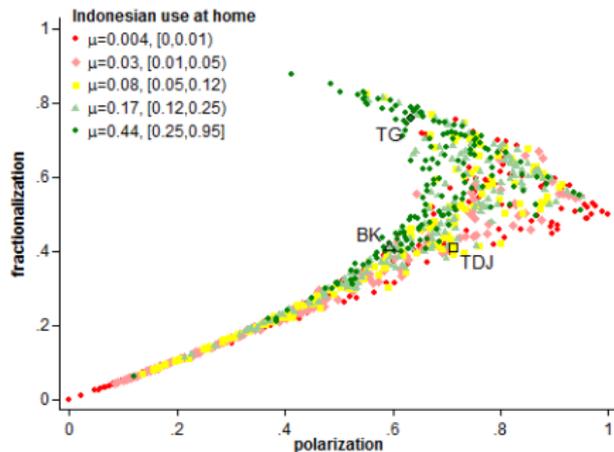
	(1)	(2)	(3)	(4)
ethnic fractionalization	1.017*** (0.095)	0.726*** (0.073)	0.599*** (0.079)	0.592*** (0.052)
ethnic polarization	-0.793*** (0.095)	-0.547*** (0.061)	-0.447*** (0.051)	-0.420*** (0.046)
Number of Villages	817	817	817	817
Number of Individuals	–	1,800,499	1,800,499	1,800,499
Dependent Variable Mean	0.145	0.154	0.152	0.152
SW fractionalization, p-value	0.000	0.000	0.000	0.000
SW polarization, p-value	0.000	0.000	0.000	0.000
KP Wald stat	7.8	8.7	10.1	22.5
Hansen J test, p-value	0.607	0.253	0.411	0.470
Hausman GMM test OLS=IV, p-value	0.372	0.807	0.747	0.769
Island FE, x Predetermined Controls	✓	✓	✓	✓
Ethnicity, Age, Relation, Gender FE			✓	✓
Birth District, Current District FE				✓

Notes: */**/** denotes significance at the 10/5/1 percent significance levels. Standard errors clustered by district.

Diversity and National Language Use At Home

A Less Parametric View

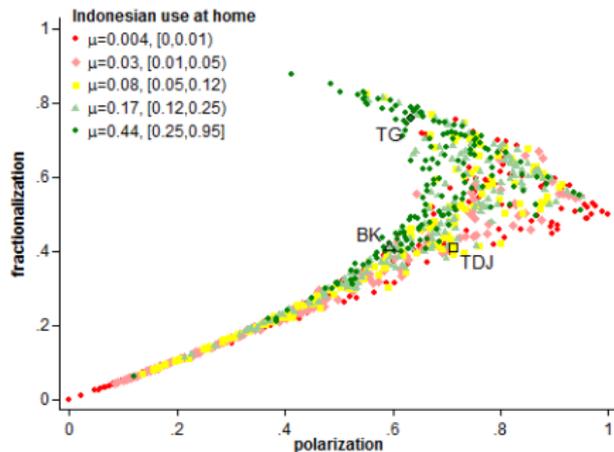
(a) Raw Data



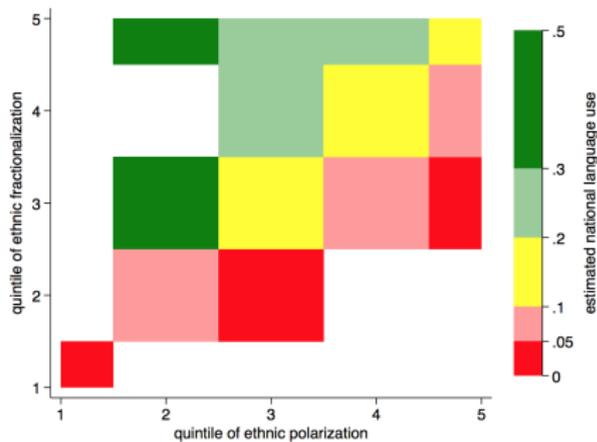
Diversity and National Language Use At Home

A Less Parametric View

(a) Raw Data

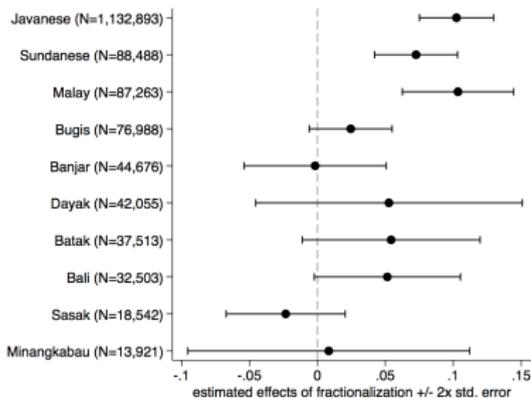


(b) Estimated Effects

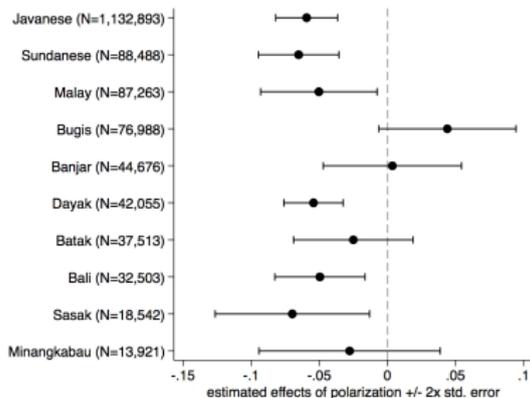


Effects on Indonesian Use at Home by Ethnicity

(a) Fractionalization



(b) Polarization



Notes: Standardized effect sizes reported. The graph reports point estimates $\pm 2 \times$ standard-error bars.

- **Malay Ethnicity:** Indonesian and Malay are mutually intelligible languages.
- 20% of Malay people in program villages report speaking Indonesian rather than Malay at home

Effects of Diversity at Different Geographic Scales

	<i>Dep. Var.: Individual Speaks National Language at Home</i>				
	(1)	(2)	(3)	(4)	(5)
ethnic fractionalization, contiguous settlements	0.054*** (0.014)				-0.006 (0.014)
ethnic fractionalization, village		0.082*** (0.011)			0.021** (0.010)
ethnic fractionalization, neighborhood			0.129*** (0.008)		0.098*** (0.009)
ethnic polarization, contiguous settlements	-0.026*** (0.009)				0.000 (0.010)
ethnic polarization, village		-0.040*** (0.008)			-0.011 (0.009)
ethnic polarization, neighborhood			-0.064*** (0.008)		-0.055*** (0.009)
2 out of 2 next-door neighbors of different ethnicity				0.192*** (0.010)	0.146*** (0.008)
1 out of 2 next-door neighbors of different ethnicity				0.067*** (0.006)	0.035*** (0.003)
Number of Villages	1,758,030	1,758,030	1,758,030	1,758,030	1,758,030
Dependent Variable Mean	0.154	0.154	0.154	0.154	0.154
R ²	0.276	0.282	0.301	0.301	0.316

Notes: The diversity measures are normalized to mean zero, standard deviation one.

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity \times current district, birth district \times current district

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity \times current district, birth district \times current district
2. restricting to villages that 'comply' with *de jure* inner-outer mix

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity × current district, birth district × current district
2. restricting to villages that 'comply' with *de jure* inner-outer mix
3. province × year-of-settlement fixed effects

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity × current district, birth district × current district
2. restricting to villages that 'comply' with *de jure* inner-outer mix
3. province × year-of-settlement fixed effects
4. similar results in other, survey data from 1995 and 2006

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity \times current district, birth district \times current district
2. restricting to villages that 'comply' with *de jure* inner-outer mix
3. province \times year-of-settlement fixed effects
4. similar results in other, survey data from 1995 and 2006
5. similar results using diversity from 2000 Pop. Census
→ also, instrumenting 2010 with 2000 diversity

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity \times current district, birth district \times current district
2. restricting to villages that 'comply' with *de jure* inner-outer mix
3. province \times year-of-settlement fixed effects
4. similar results in other, survey data from 1995 and 2006
5. similar results using diversity from 2000 Pop. Census
→ also, instrumenting 2010 with 2000 diversity
6. other measures of ethnic diversity, e.g., aggregated ethnic groups

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity \times current district, birth district \times current district
2. restricting to villages that 'comply' with *de jure* inner-outer mix
3. province \times year-of-settlement fixed effects
4. similar results in other, survey data from 1995 and 2006
5. similar results using diversity from 2000 Pop. Census
→ also, instrumenting 2010 with 2000 diversity
6. other measures of ethnic diversity, e.g., aggregated ethnic groups
7. null effects of religious fractionalization or polarization

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity \times current district, birth district \times current district
2. restricting to villages that 'comply' with *de jure* inner-outer mix
3. province \times year-of-settlement fixed effects
4. similar results in other, survey data from 1995 and 2006
5. similar results using diversity from 2000 Pop. Census
→ also, instrumenting 2010 with 2000 diversity
6. other measures of ethnic diversity, e.g., aggregated ethnic groups
7. null effects of religious fractionalization or polarization
8. other village-level controls not explicitly used to select sites

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity \times current district, birth district \times current district
2. restricting to villages that 'comply' with *de jure* inner-outer mix
3. province \times year-of-settlement fixed effects
4. similar results in other, survey data from 1995 and 2006
5. similar results using diversity from 2000 Pop. Census
→ also, instrumenting 2010 with 2000 diversity
6. other measures of ethnic diversity, e.g., aggregated ethnic groups
7. null effects of religious fractionalization or polarization
8. other village-level controls not explicitly used to select sites
9. diversity beyond own-village in contiguous Transmigration settlements

Robustness

1. different fixed effects rule out several confounders
→ others: ethnicity × current district, birth district × current district
2. restricting to villages that 'comply' with *de jure* inner-outer mix
3. province × year-of-settlement fixed effects
4. similar results in other, survey data from 1995 and 2006
5. similar results using diversity from 2000 Pop. Census
→ also, instrumenting 2010 with 2000 diversity
6. other measures of ethnic diversity, e.g., aggregated ethnic groups
7. null effects of religious fractionalization or polarization
8. other village-level controls not explicitly used to select sites
9. diversity beyond own-village in contiguous Transmigration settlements
10. results hold within many subgroups in the data
→ children with parents who (i) do not speak Indonesian at home
(ii) are interethnically married or not
→ education, occupation, inner-born, outer-born . . .

Addressing Sorting by Sub-Populations

	<i>Dep. Var.: Individual Speaks National Language at Home</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	baseline	inner-ethnic	inner-born < yr. settled	outer-ethnic	outer-born < yr. settled	outer-born APPDT	outer-born non-APPDT	born same distr. ≥ yr. settled
ethnic fractionalization	0.082*** (0.011)	0.098*** (0.013)	0.081*** (0.012)	0.056*** (0.015)	0.069*** (0.015)	0.069*** (0.020)	0.056*** (0.015)	0.098*** (0.014)
ethnic polarization	-0.040*** (0.008)	-0.058*** (0.010)	-0.053*** (0.010)	-0.028** (0.012)	-0.024** (0.011)	-0.035*** (0.012)	0.001 (0.014)	-0.046*** (0.011)
Number of Individuals	1,800,499	1,267,946	543,655	532,486	408,751	282,030	126,721	626,772
Dependent Variable Mean	0.154	0.099	0.066	0.285	0.207	0.158	0.316	0.168
R ²	0.281	0.198	0.143	0.328	0.299	0.305	0.283	0.300
Island FE, x Controls	✓	✓	✓	✓	✓	✓	✓	✓
Ethnicity, Age, Gender FE	✓	✓	✓	✓	✓	✓	✓	✓

Addressing a Purely Economic Interpretation

Similar Effects across Education Levels

	baseline	no school	primary		junior	secondary	
	(1)	(2)	some	completed	(5)	senior	post-
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ethnic fractionalization	0.082*** (0.011)	0.057*** (0.009)	0.082*** (0.012)	0.072*** (0.010)	0.088*** (0.013)	0.095*** (0.013)	0.057*** (0.016)
ethnic polarization	-0.040*** (0.008)	-0.029*** (0.007)	-0.036*** (0.010)	-0.042*** (0.007)	-0.042*** (0.010)	-0.028** (0.013)	-0.006 (0.014)
Number of Individuals	1,800,499	141,545	408,269	650,912	336,498	198,334	64,070
Dependent Variable Mean	0.154	0.116	0.165	0.102	0.156	0.260	0.347
R ²	0.281	0.324	0.308	0.250	0.276	0.294	0.304

Addressing a Purely Economic Interpretation

Similar Effects across Occupations

	baseline (1)	not working (2)	agri/mine (3)	manuf. (4)	manual (5)	trade/svc (6)	white collar (7)	other (8)
ethnic fractionalization	0.080*** (0.011)	0.089*** (0.013)	0.058*** (0.008)	0.075*** (0.016)	0.107*** (0.015)	0.081*** (0.012)	0.071*** (0.016)	0.092*** (0.017)
ethnic polarization	-0.041*** (0.008)	-0.042*** (0.010)	-0.034*** (0.007)	-0.026** (0.012)	-0.057*** (0.014)	-0.035*** (0.011)	-0.018 (0.015)	-0.028* (0.015)
Number of Individuals	1,590,709	685,523	640,488	21,372	27,246	97,930	87,272	10,374
Dependent Variable Mean	0.143	0.165	0.085	0.163	0.152	0.191	0.305	0.205
R ²	0.276	0.286	0.241	0.336	0.327	0.280	0.313	0.325

Potential Mechanisms: Intergroup Distances

1. **Spatial:** residential ethnic segregation
2. **Economic:** interethnic inequality
3. **Cultural:** ethnolinguistic distance

1. Spatial Distribution of Ethnic Groups

- Measure village segregation using Alesina & Zhuravskaya (2011), which generalizes 2-group segregation indices to many groups

$$segregation_v = \frac{1}{G-1} \sum_{g=1}^G \sum_{b=1}^B \left(\frac{n_{bv}}{N_v} \right) \frac{(\pi_{bgv} - \pi_{gv})^2}{\pi_{gv}}$$

where ethnic groups $g \in G$ and census blocks $b \in B$

- Recall that **initial housing** and farm plots **allocated by lottery**
- For identical levels of diversity (F, P), segregation is significantly lower in Transmigration than non-Transmigration villages [▶ table](#)
- We exploit this **policy-induced** variation to isolate **local contact**
→ instrument overall segregation w/ segregation of initial, old settlers

1. Spatial Distribution of Ethnic Groups

Segregation Dampens Competing Effects of Diversity

	<i>Dep. Var.:</i> Share of Village using Indonesian at Home	
	(1)	(2)
fractionalization	0.135*** (0.015)	0.145*** (0.015)
polarization	-0.068*** (0.011)	-0.084*** (0.012)
segregation	-0.033*** (0.006)	-0.031*** (0.006)
fractionalization \times segregation		-0.041*** (0.006)
polarization \times segregation		0.018** (0.007)
Number of Villages	817	817

2. Interethnic Inequality

- Predetermined interethnic inequality may exist
- This could influence the nature of intergroup contact
- We measure inequality using a typical Greenberg-Gini formulation

$$\text{interethnic inequality}_v = \frac{1}{2\bar{x}_v} \sum_{i=1}^I \sum_{j=1}^J n_{iv} n_{jv} |\bar{x}_{iv} - \bar{x}_{jv}|$$

- x : agroclimatic similarity b/t individuals' origin and destination
→ key determinant of productivity in new settlements (Bazzi et al, 2016)

2. Interethnic Inequality

Inequality Reduces Positive Effects of F

	<i>Dep. Var.: Share Village Indonesian at Home</i>	
	(1)	(2)
fractionalization	0.149*** (0.015)	0.165*** (0.019)
polarization	-0.073*** (0.012)	-0.102*** (0.016)
interethnic inequality, agroclimatic similarity	-0.033*** (0.009)	-0.019 (0.012)
fractionalization \times interethnic inequality		-0.036** (0.015)
polarization \times interethnic inequality		0.012 (0.011)
Number of Villages	817	817
Dep. Var. Mean	0.145	0.145

Notes: All measures are normalized to mean zero, standard deviation one. Also controls for mean agroclimatic similarity.

3. Predetermined Cultural Distance

- Predetermined cultural differences may amplify effects of diversity
- We consider diversity metrics incorporating cultural distances:

$$\text{fractionalization}(\delta) = \sum_{i=1}^I \sum_{j=1}^J n_i n_j \delta_{ij}$$

$$\text{polarization}(\delta) = \sum_{i=1}^I \sum_{j=1}^J n_i^2 n_j \delta_{ij}$$

where δ_{ij} captures predetermined linguistic distance b/t groups i and j

- We parametrize δ_{ij} following Fearon (2003)

$$\delta_{ij} = 1 - \left(\frac{\text{shared language tree branches}_{ij}}{\text{total possible shared branches}} \right)^\tau$$

where $\tau = 0.05$ as in Esteban et al (2012) to capture deeper cleavages

3. Predetermined Cultural Distance

Adjusted F and P Have Larger Effect Sizes

	<i>Dep. Var.: Shr. Village Indonesian at Home</i>	
	(1)	(2)
fractionalization	0.135*** (0.015)	
polarization	-0.083*** (0.012)	
fractionalization(δ)		0.144*** (0.016)
polarization (δ)		-0.092*** (0.013)
Number of Villages	817	817
Dep. Var. Mean	0.145	0.145
$H_0 : F(\delta) = F$ baseline, p-value		[0.100]
$H_0 : P(\delta) = P$ baseline, p-value		[0.022]**

Notes: All measures are normalized to mean zero, standard deviation one.

Broader Evidence of Nation Building

1. **Intergroup tolerance and trust, public goods contribution**
(*Susen*s 2012 household survey, covers 87 Transmigration villages)
2. **Interethnic marriage rates, adjusted for potential rates**
(similar results using 2000 and 2010 Pop. Census)
3. **Names of children born from 2000–2010**
(2010 Pop. Census using prior diversity from 2000 Pop. Census)
4. **Ethnic conflict, public goods, development**
(*SNPK*, NOAA, *Susen*s, and various rounds of *Podes*)

1. Intergroup Preferences, Trust, and Public Goods

<i>Dependent Variable</i>	fractionalization (F)	polarization (P)	Dep. Var. μ (1-5 scale)
1. voluntary public good provision	0.166 (0.113)	-0.224* (0.119)	2.7
2. join community group(s)	0.017 (0.129)	-0.068 (0.106)	2.4
3. pleased with non-coethnics	0.106 (0.189)	-0.285* (0.167)	2.9
4. trust neighbor to watch house	0.145 (0.120)	-0.242** (0.100)	2.9
5. trust neighbor to tend children	-0.080 (0.149)	-0.120 (0.124)	2.7
6. feel safe	-0.077 (0.107)	-0.202** (0.099)	3.2
7. ease in obtaining neighbor assistance	0.005 (0.121)	-0.120 (0.104)	2.7
8. contribute to assist unfortunate neighbors	0.227** (0.097)	-0.199* (0.113)	2.9

Notes: At most 860 individuals in 87 Transmigration villages. Controls for gender, age and education.

n.b.: including only F delivers negative results consistent with prior literature

2. Interethnic Marriage

Among Young Individuals Married Post-Settlement

	Post-Settlement Inter-marriage Rate in			
	2000 actual (1)	2010 (2)	2000 supply-adjusted (3)	2010 (4)
ethnic fractionalization	0.068*** (0.012)	0.093*** (0.008)	-0.025 (0.022)	-0.006 (0.013)
ethnic polarization	-0.028*** (0.010)	-0.027*** (0.007)	-0.081*** (0.021)	-0.112*** (0.012)
Number of Villages	815	817	815	816
Dependent Variable Mean	0.152	0.178	0.388	0.482
R ²	0.258	0.562	0.114	0.317

Notes: Supply-adjusted divides the actual intermarriage rate by the rate implied by random matching within relevant age range.

3. Children's Name Choices

- We classify the content of **names** to capture **identity**
 1. Indonesian language use at home
 2. intermarried
 3. urban resident
 4. own-group ethnicity
- Suppose you are trying to guess child's identity group g (1)–(4)

3. Children's Name Choices

- We classify the content of **names** to capture **identity**
 1. Indonesian language use at home
 2. intermarried
 3. urban resident
 4. own-group ethnicity
- Suppose you are trying to guess child's identity group g (1)–(4)
- Name precision \rightarrow positive predictive value, stronger identity

$$\begin{aligned} PRE(\text{name}|g) &= \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \\ &= \frac{Pr\{\text{name}_i = n \mid g_i = g\}}{Pr\{\text{name}_i = n \mid g_i = g\} + Pr\{\text{name}_i = n \mid g_i \neq g\}} \end{aligned}$$

where target pop. is outside of Transmigration villages (> 10 km)

- Only calculate this for names shared by 100 or more people.
- Restrict to children born July 2000–2010 in the 2010 Pop. Census

3. Diversity and Children's Name Choices

	<i>Dep. Var.:</i> precision of name in identifying ...			
	Indonesian language home (1)	intermarried household (2)	urban (3)	own-ethnicity (4)
ethnic fractionalization	0.222*** (0.038)	0.196*** (0.041)	0.268*** (0.052)	-0.215*** (0.042)
ethnic polarization	-0.127*** (0.032)	-0.113*** (0.034)	-0.161*** (0.044)	0.081** (0.032)
Number of Individuals	726,969	676,307	731,628	720,142
R ²	0.101	0.190	0.080	0.101

Notes: Standardized outcomes. Restricted to children in 2010 Census born after the 2000 Census round. Diversity measures based on 2000 Census. Includes ethnicity and age fixed effects. Normalized outcome measures based on the *PRE* index from Fryer & Levitt (2004). Similar results using double-metaphone transformation of reported names.

4. Conflict, Public Goods and Development

	local development and public goods			conflict		voting	
	village pub. goods (1)	light intensity (2)	household exp./capita (3)	any ethnic conflict <i>SNPK</i> (4)	<i>Podes</i> (5)	turnout (6)	<i>Pancasila</i> party 1st-3rd (7)
ethnic fractionalization	0.030** (0.011)	0.026* (0.015)	0.067** (0.033)	-0.062** (0.028)	-0.005 (0.004)	-0.001 (0.006)	-0.022 (0.032)
ethnic polarization	-0.022* (0.011)	-0.025* (0.014)	-0.038 (0.036)	0.066** (0.028)	0.004 (0.004)	-0.003 (0.007)	-0.045 (0.031)
Number of Villages	817	817	710	244	817	795	817
Dependent Variable Mean	0.412	0.082	12.489	0.045	0.010	0.947	0.470
R ²	0.227	0.109	0.124	0.316	0.028	0.092	0.106

Notes: Diversity measured in 2000. Outcomes measured from 2000 onward.

Intergenerational Nation Building Process

Taking Stock

1. Parents arrive with children in new settlement
(initial diversity and segregation determined exogenously)
 2. Gradually, make investments in children's identity
(language use at home)
 3. Children go on to form new households
(intermarriage, language use at home, names for their kids)
 4. Intergroup preferences, trust, and public goods update throughout
- consistent with dynamic model of growth in national identity ...

Roadmap

Introduction

Background: Diversity and the Transmigration Program

Data: Diversity and Nation Building Outcomes

Results: Diversity, Socialization, and Identity

- National Language Use at Home

- Mechanisms

- Other Outcomes

A Model of Growth in National Identity

Conclusion

Model Snapshot

- We model identity dynamics using evolutionary game theory
- Inspired by Darity et al. (2006) on racist whites and blacks
- Delivers Δ national identity through social interaction amid diversity
- Three key features driving growth in national identity
 1. Interethnic trade benefit
 2. Relative costs of switching from ethnic to national identity
 3. Disutility from interethnic cultural competition
- **Theoretical Results:**
 1. Equilibrium Selection: F increases basin of attraction for national identity, P shrinks basin of attraction
 2. Replicator Dynamics: F hastens growth of national identity, P slows growth

Setup

- **Ethnic groups:** indexed by $j, k = 1, \dots, J$ fixed and exogenous
- Each ethnic group has two **strategies** (actions):
 1. **nationalist:** N ; share: π_k
 2. **ethnic loyal:** E ; share: $(1 - \pi_k)$
- In each period, agents interact through matching
→ **payoffs:** gains from trade, costs of failure to cooperate
- Assume pure random matching (given lottery-assigned housing):

$$\mathbb{P}(\text{matched to group } k) = \left(\frac{\text{population}_k}{\text{population}} \right) = p_k$$

$$\mathbb{P}(\text{matched to type } N \text{ from group } k) = p_k \pi_k$$

(can also make this a function of segregation σ)

Benefits and Costs of Identity

1. Gains from Non-Market Interactions: θ

- Only if you are from same group, or share the same identity

2. Identity Costs:

- γ_E : Ethnic-loyal identity cost (each period)
- γ_N : National identity cost (each period)
- set $\gamma \equiv \gamma_N - \gamma_E$: differential identity cost

3. Coordination Failure: δ

- Relative penalty that type N gets from matching w/ type E from another group
- Set $\delta = \psi p_k$, i.e., increases in other group size
- Interpretation: interethnic antagonism, loss in “protection” from own group (club good model)

Payoffs

		Member of Group j	
		Ethnic-loyal j	Nationalist j
Matched With	Ethnic-loyal j	$\theta - \gamma_E$	$\theta - \gamma_N$
	Nationalist j	$\theta - \gamma_E$	$\theta - \gamma_N$
	Ethnic-loyal k	$-\gamma_E - S_k^E$	$-\gamma_N - S_k$
	Nationalist k	$-\gamma_E$	$\theta - \gamma_N$

- We set $\delta_k \equiv S_k - S_k^E = \psi p_k$ with $\psi > 0$

Remarks

- Players are assigned strategies but do not choose them (“biological model” for the evolution of identity choices)
- Fitness of given strategy reflected in expected payoffs
- Players imitate successful strategies
- Fitter strategies become more prevalent over time

Replicator Dynamic

- **Growth Rate of the National Identity** strategy for group j :

$$\frac{\partial \pi_j}{\partial t} = \pi_j (1 - \pi_j) \left\{ \underbrace{\theta \sum_{k \neq j} p_k \pi_k}_{\text{relative gain from trade}} - \underbrace{\psi \sum_{k \neq j} (1 - \pi_k) p_k^2}_{\text{relative interethnic antagonism}} - \underbrace{\gamma}_{\text{relative identity cost}} \right\}$$

- **Trade Benefit:** when this is larger, incentives for N increase
 - Weight depends on share of nationals in other groups (Lazear, 1999)
- **Relative Identity Cost:** reflects difference in adoption costs
- **Disutility from Coordination Failure:** \uparrow square of other group shr.

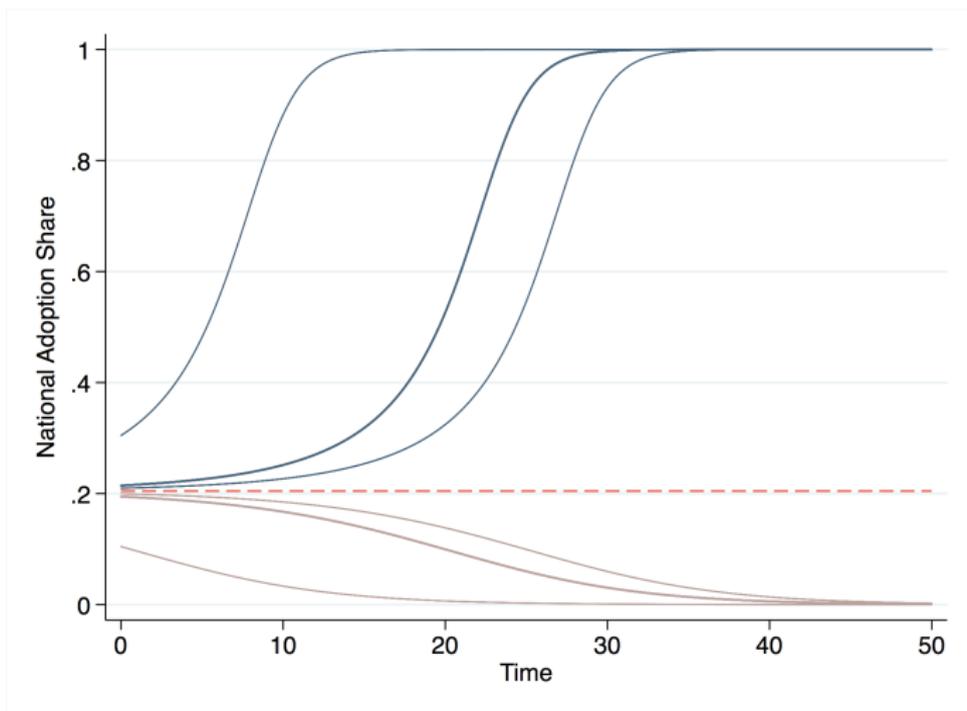
Replicator Dynamic

- **Growth Rate of the National Identity** strategy for group j :

$$\frac{\partial \pi_j}{\partial t} = \pi_j (1 - \pi_j) \left\{ \underbrace{\theta \sum_{k \neq j} p_k \pi_k}_{\text{relative gain from trade}} - \underbrace{\psi \sum_{k \neq j} (1 - \pi_k) p_k^2}_{\text{relative interethnic antagonism}} - \underbrace{\gamma}_{\text{relative identity cost}} \right\}$$

- **Trade Benefit:** when this is larger, incentives for N increase
 - Weight depends on share of nationals in other groups (Lazear, 1999)
- **Relative Identity Cost:** reflects difference in adoption costs
- **Disutility from Coordination Failure:** \uparrow square of other group shr.
- **Approximation Argument:** we show that:
 - $\partial \pi_j / \partial t$ increases as F increases
 - $\partial \pi_j / \partial t$ decreases as P increases

Multiple Evolutionary Equilibria and Basins of Attraction



- We solve for **tipping points** and show that:
 - Increasing F increases basin of attraction to national identity
 - Increasing P reduces basin of attraction to national identity

Roadmap

Introduction

Background: Diversity and the Transmigration Program

Data: Diversity and Nation Building Outcomes

Results: Diversity, Socialization, and Identity

- National Language Use at Home

- Mechanisms

- Other Outcomes

A Model of Growth in National Identity

Conclusion

Key Takeaways

A Policy Experiment in Ethnic Mixing

- Different dimensions of diversity have different impacts
- Ethnic Fractionalization \implies integration, less ethnic attachment
 - Consistent w/ nation building
- Ethnic Polarization \implies increased ethnic attachment, entrenchment
- Segregation and ethnic inequality reverse effects of diversity
- Theory: polarization and fractionalization shape identity formation
- Policy: lessons on how to manage diversity

External Validity and Broader Relevance

Public Policy around Intergroup Cooperation

- Migration and **resettlement** pressures rising globally
 - Resettlement policy challenge due to conflict, climate change, etc (de Sherbenin et al, 2011)
- **Integration policies** in OECD countries w/ growing immigration
- Any role for **state-sponsored internal migration** given many examples outside Indonesia of less benign intentions and outcomes?
 - growing evidence: spontaneous migration \implies 'sons of the soil' conflict
- **Language policy**: national vs. official vs. majority
 - National language in Indonesia compared to India and Philippines
 - Success of Swahili in Tanzania
 - French as unifying language in historical France
 - Ongoing debates in Spain, Sri Lanka, . . .

THANKS!

Email: agaduh@walton.uark.edu

APPENDIX

Examples of Inner–Outer Group Differences

based on interethnic marriages observed in study villages

1. Javanese vs. Batak (North Sumatra)

- Batak have bride price while Javanese typically do not
- Batak are patrilocal while Javanese are matrilocal
- Batak have patrilineal inheritance while Javanese have equal inheritance

2. Javanese vs. Minang (West Sumatra)

- Minang and Javanese have no marital wealth exchange traditions
- Minang have no common post-marital residence rules while Javanese are matrilocal
- Minang have matrilineal inheritance while Javanese have equal inheritance

3. Balinese vs. Toraja (Central Sulawesi)

- Toraja have bride price while Balinese typically do not
- Toraja are matrilocal while Balinese are patrilocal
- Toraja have equal inheritance while Balinese have patrilineal inheritance

Generally, Inner–Inner differences dwarfed by Inner–Outer differences

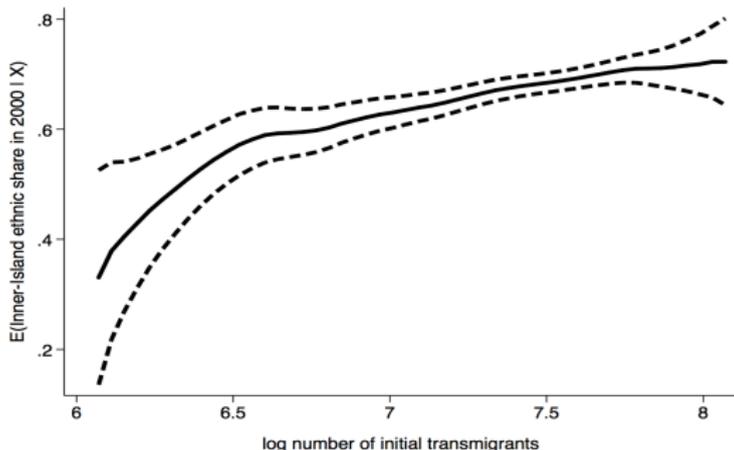
Examples of Inner–Outer Group Differences based on interethnic marriages observed in study villages

Linguistic Differences

Language	Branches						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Javanese	Austronesian	Malayo-Polynesian	Javanese				
Minangkabau	Austronesian	Malayo-Polynesian	Malayo-Sumbawan	North and East	Malayic	Malay	
Batak	Austronesian	Malayo-Polynesian	NW Sumatra-Barrier Islands	Batak	Southern		
Toraja	Austronesian	Malayo-Polynesian	South Sulawesi	Northern	Toraja-Sa'dan		
English	Indo-European	Germanic	West	English			
German	Indo-European	Germanic	West	High German	German	Middle German	East Middle German
French	Indo-European	Italic	Romance	Italo-Western	Western	Gallo-Iberian	Gallo-Romance

Notes: Ethnologue language classification.

Instrument Quality



Several results support **excludability**. IV uncorrelated with:

1. linguistic distance b/t Inner- and indigenous Outer-Island ethnicity
2. ex post immigration (by group) between 1995 and 2000
3. agroclimatic similarity of transmigrants (proxy for economic welfare)
4. other measures of diversity (e.g., birthplace or religious diversity)
5. other measures of predetermined local political and economic development (not explicitly used by the planners)

Diversity and Language Use at Home

	<i>Dep. Var.: [...] is Main Language at Home</i>					
	Indonesian		Native Ethnic		Other Ethnic	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
fractionalization	0.359*** (0.049)	0.401*** (0.108)	-0.351*** (0.043)	-0.450*** (0.036)	-0.007 (0.040)	0.060** (0.023)
polarization	-0.173*** (0.038)	-0.141*** (0.022)	0.180*** (0.033)	0.245*** (0.040)	-0.007 (0.027)	-0.083*** (0.028)
Number of Individuals (\geq age 5)	1,800,499	1,799,160	1,800,310	1,799,160	1,800,310	1,799,160
Dep. Var. Mean	0.154	0.154	0.764	0.764	0.082	0.082

Notes: Full fixed effects specification. Sanderson & Windmeijer (2016) weak-instrument test p-value < 0.001 for both fractionalization and polarization.

Residential Segregation is Significantly Lower in Transmigration Program Villages than Non-Program Villages

	Villages >10km from Transmigration Villages		"Almost-Treated" Control Villages	
	(1)	(2)	(3)	(4)
Transmigration village	-0.006*** (0.002)	-0.004*** (0.002)	-0.012*** (0.004)	-0.010*** (0.003)
Number of Villages	23,562	23,562	1,514	1,514
Dependent Variable Mean	0.020	0.020	0.029	0.029
R ²	0.262	0.305	0.225	0.383

Notes: */**/** denotes significance at the 10/5/1 percent significance levels.

Probing Instrument Validity

- Our instruments—initial number of Java/Bali-born transmigrants and ethnic diversity among them—are uncorrelated with:
 1. linguistic distance b/t Inner and indigenous Outer-Island ethnicity
 2. ex post immigration (by group) between 1995 and 2000
 3. agroclimatic similarity of transmigrants (proxy for economic welfare)
 4. other measures of diversity (e.g., birthplace or religious diversity)
 5. other measures of predetermined local political and economic development not explicitly used by the planners