Modernization, Social Identity, and Ethnic Conflict

Kazuhiro Yuki*

This version: July 2020
First version: March 2015

Abstract

Evidence suggests that ethnic divisions in a society lead to negative outcomes in conflict and development. It is often argued that the lack of shared national identity lies behind the poor performance. If shared identity is important, how can it be realized? Conflicting theses exist in political science regarding the effect of the modernization (e.g., industrialization and urbanization) of a society on national identity; the classic thesis argues the positive effect, while the competing thesis argues the negative effect, suggesting that policies promoting modernization have negative effects on national identity and conflict and thus might not be very effective for development. Which thesis is more relevant under what conditions? Meanwhile, some scholars stress the effectiveness of nation-building policies in strengthening national identity. Do nation-building policies influence the consequences of modernization for identity, conflict, and development?

To examine these questions theoretically, a model of social identity, ethnic conflict, and development is developed. It is shown that, as modernization proceeds, a society shifts to an equilibrium with a universal national identity and good results in conflict and development, if national pride is high, ethnic differences are not salient in people’s minds, contested resources are not abundant, or institutions are good in quality; otherwise, it shifts to an equilibrium with a universal ethnic identity and worse outcomes in other dimensions. Hence, the result is consistent with the classic (competing) thesis on the effects of modernization on identity under the former (latter) situation. Analysis suggests that nation-building policies play a role in overcoming the negative effects of modernization under the latter situation.

Keywords: ethnic conflict, social identity, modernization, nation building, economic development

JEL classification numbers: D72, D74, O10, O20

*Faculty of Economics, Kyoto University, Yoshida-hommachi, Sakyo-ku, Kyoto, 606-8501, Japan; Phone +81-75-753-3532; E-mail yuki@econ.kyoto-u.ac.jp. Invaluable comments and suggestions from two anonymous referees are gratefully appreciated. Helpful comments are also provided by participants at the ERF workshop on Macro-economics and the 2016 Asian Meeting of the Econometric Society. Financial support from JSPS through Grants-in-Aid for Scientific Research 10197395 is acknowledged.
1 Introduction

Empirical evidence suggests that ethnic divisions in a society lead to negative outcomes in various dimensions, including civil conflict (Esteban, Mayoral, and Ray, 2012) and economic development (Montalvo and Reynal-Quero, 2005). It is often argued that the lack of shared social identity, that is, the dominance of subnational (particularly, ethnic) identities over national identity, lies behind the negative outcomes in ethnically heterogeneous societies (Collier, 2009; Michalopoulos and Papaioannou, 2015).

If shared national identity is important, how can it be realized? Does the modernization (e.g., industrialization, the diffusion of education, and urbanization) of a society bring shared identity? In political science, competing theses exist regarding the effect of modernization on national identity (Robinson, 2014). The classic thesis, based on the past experience of Europe, argues that modernization leads to widespread national identity at the expense of ethnic and other subnational identities (Deutsch, 1953; Gellner, 1964, 1983; Weber, 1979). Based on the post-independent experience of Africa, another influential thesis asserts that modernization breeds ethnic identification due to intensified competition over resources (Melson and Wolpe, 1970; Bates, 1983). If the classic view is correct, policies promoting modernization might be sufficient in attaining shared national identity and good results in conflict and development, whereas if the competing view is true, such policies have negative effects on national identity and conflict and thus might not be very effective for development. Which view is more relevant under what conditions?

Meanwhile, Miguel (2004), Collier (2009), and Blouin and Mukand (2019), based on a case study or statistical analysis, argue that nation-building policies, such as school education or government propaganda emphasizing common history, culture, and values and the promotion of a national language, are effective in strengthening national identity.¹ Do these policies influence the consequences of modernization for identity, conflict, and development?

To examine these questions theoretically, this paper develops a model of social identity, ethnic conflict, and economic development. In the model, which builds on the model of social identification and ethnic conflict by Sambanis and Shayo (2013), individuals choose a sector to work for (between the modern sector and a traditional sector), a social identity (between ethnic and national identities), and a contribution to conflict. The degree of modernization (and output), identity, and conflict interact each other.

The seminal model by Sambanis and Shayo (2013) is explained first to articulate main features of the model of this paper. In a standard contest model of conflict in which different groups compete for exogenous resources used to produce group-specific club goods (e.g., public services and infrastructures benefiting a specific group), members of each group decide how much they

¹Miguel (2004) finds that two neighboring rural districts of Tanzania and Kenya, which largely shared geography, history, and a colonial institutional legacy, exhibit a sharp difference in the relationship between ethnic diversity and the local provision of public goods (school funds and infrastructures), significantly negative for the Kenyan district and positive but insignificant for the Tanzanian district. He also finds that the relationship is insignificant for other local public finance outcomes for Tanzania (no comparable data for Kenya). He argues that sharply different ethnic policies in areas such as national language and public school education by post-independent governments contributed to differences in the strength of national identity and the above-mentioned relationship between the two countries.
contribute to conflict taking into account individual cost and benefit of the activity, and the total
contribution of group members determine the amount of resources the group obtains. Sambanis
and Shayo (2013) augment the standard model with socio-psychological factors. Individual utility
depends not only on individual cost and benefit of contributing to conflict but also negatively on
the perceived distance to a social group the individual identifies with (his or her ethnic group or
the nation) and positively on (iii) the status of the social group. That is, an individual suffers a
large cognitive cost when he or she is very different from others of the group in relevant aspects,
while the individual takes pride in being a member of the group when the status of the group is
high. These socio-psychological components are important factors affecting social identification
and intergroup behaviors, according to influential theories in social psychology (Tajfel and Turner,
1986; Turner et al., 1987) and empirical evidence (Manning and Roy, 2010; Sambanis, Skaperdas,
and Wohlforth, 2015). Importantly, because the socio-psychological components differ depending
on which group an individual identifies with, individual contribution to ethnic conflict and thus the
aggregate level of conflict are influenced by social identities of individuals. Further, the choice of
social identity is endogenous: individuals choose to identify with a group that brings them higher
utilities. Hence, social identity and individual and aggregate outcomes interact each other.

In order to examine the effect of modernization on social identity, conflict, and output, the
present model modifies the Sambanis and Shayo model on the following points. First, there are
multiple sectors producing private goods — ethnically segregated traditional sectors, which repre-
sents sectors such as traditional agriculture, the urban informal sector, and household production
in the real economy, and the ethnically integrated modern sector — and individuals choose the
sector maximizing their utilities. The modernization of a society is interpreted as the sectoral shift
of labor and output from the traditional sectors to the modern sector and is driven by the increased
productivity of the modern sector. Second, sectoral affiliation, along with ethnicity, is a component
of the perceived distance. That is, when measuring how similar to or different from other members
of a social group (an ethnic group or the nation), individuals take into account sectoral affiliations
as well as ethnicity of themselves and other members.

Although the present model is a simple extension of the Sambanis and Shayo model, it can
examine not only the interaction between social identity and conflict, on which Sambanis and
Shayo (2013) focus, but also the interactions of social identity with economic variables, i.e., sectoral
compositions of labor and output and aggregate output. By examining how productivity-driven
modernization affects these interactions, this paper provides answers to an important question
long discussed in political science (How does the modernization of a society affect social identity
and conflict?) and related questions (Are policies promoting modernization always effective for

---

2 The concept of perceived distance is the basis of a major social psychological theory, self-categorization theory
(Turner et al., 1987). Intergroup status differences are major factors affecting intergroup behaviors such as conflict
and discrimination, according to a closely-related influential theory, social identity theory (Tajfel and Turner, 1986).
3 Evidence suggests that perceived distance and status affect identity. Manning and Roy (2010) find for the UK
that nonwhites, whose perceived distance to the nation seems to be greater, are less likely to think of themselves as
British than whites. Further, they find that immigrants from poorer and less democratic (i.e., lower status) countries
assimilate faster into British identity. See footnote 14 for evidence from Sambanis, Skaperdas, and Wohlforth (2015).
conflict reduction and development? Do nation-building policies influence the consequences of modernization for identity and other outcomes?) This is the main contribution of the paper.

For analytical tractability and focus on an ethnically diverse society without a dominant ethnic group, the paper assumes that ethnic groups are symmetric in every aspect and examines equilibria in which choices of all groups are symmetric. Such symmetric equilibria can be classified into two types: equilibria in which individuals of the same ethnic group share the same identity (homogenous identity equilibria) and equilibria in which they do not (heterogenous identity equilibria). In every equilibrium, modern sector workers are more (less) likely to identify with the nation (their ethnic group) than traditional sector workers, which is consistent with empirical evidence for African nations (Robinson, 2014).

Firstly, different equilibria are compared for given parameters and exogenous variables. (Multiple or even all equilibria exist for certain parameters and exogenous variables.) Then, it is found that, in an equilibrium with a higher proportion of individuals identifying with the nation, the level of conflict is lower, the shares of modern sector workers and output are higher, and under plausible conditions, the total output of private goods and aggregate material payoff (the value of private and club goods consumption net of the cost of conflict) are higher. That is, national identity is associated with not only a lower level of ethnic conflict, which is shown in Sambanis and Shayo (2013), but also higher modern sector shares and higher output.4

Whereas the economic and political outcomes differ depending on which equilibrium is realized for given parameters and exogenous variables, the set of realized equilibria and the outcomes change with values of exogenous variables. In particular, by examining how they change with increased (total factor) productivity of the modern sector, the paper analyzes the effect of modernization on identity, conflict, and output. The productivity growth raises the modern sector wage, induces a higher proportion of workers to choose the sector, and raises the sector’s share in production. How does productivity-driven modernization affect social identity, conflict, and aggregate output?

The answer depends on the difference between the status of the nation and the status of ethnic groups, which is denoted the status difference. The national status and thus the status difference would be high when the people of a nation believe that they share glorious history, rich culture, or a "right" sense of values, because they would feel proud of belonging to such nation, while the ethnic status would be high and thus the status difference would be low when ethnic groups are distinctive in these dimensions.

If the status difference is at extreme levels, the society stays in the same equilibrium: when the national status compared with the ethnic status is very high (very low), everyone always identifies with the nation (his or her ethnic group) and the level of conflict is consistently low (high). Otherwise, when the status difference is relatively high (low) and the society starts with a heterogenous identity equilibrium, the society shifts from the equilibrium, in which modern sector

4The lower share of modern sector workers under ethnic identity indicates the greater intersectoral gap in earnings. This suggests that strong ethnic identity might partly explain a substantial gap in average labor productivity between agriculture and non-agriculture in many developing countries found by Gollin, Lagakos, and Waugh (2014) and others.
workers are more likely to have a national identity than traditional sector workers, to the one with a universal national (ethnic) identity and a low (high) level of conflict. That is, the social identity initially associated with modern (traditional) sector workers eventually becomes the shared identity, when the status difference is high (low). Although increased productivity always raises modern sector’s shares in employment and production, given the productivity level, the society with high (low) status difference achieves relatively large (small) modern sector shares and, under plausible conditions, relatively high (low) levels of aggregate output and material payoff. Hence, having sufficiently high national status relative to ethnic status is crucial for achieving good outcomes in development as well as in national identity and conflict.

However, history or "luck" is also important, as long as the status difference is not at extreme levels. Given parameters and exogenous variables including status, multiple equilibria exist, particularly when the level of development is not high, thus identity, conflict, and output differ depending on which equilibrium is realized. If an equilibrium realized in the initial period happens to be such that a relatively high proportion of individuals identify with the nation, the society tends to subsequently maintain a relatively strong national identity and relatively good outcomes in other dimensions.

Similar results also hold for contested resources when the "low (high) status difference" of the above result is replaced with a "large (small) amount of contested resources". Specifically, given the status difference, when the amount of contested resources is large (small), the society shifts from a heterogenous identity equilibrium to the equilibrium of universal ethnic (national) identity, or stays in the latter equilibrium. Contested resources represent not only material resources (such as natural resources) but also a part of the governmental budget that can be used to benefit particular ethnic groups, whose allocation among ethnic groups is determined not by rules but by the consequences of violent or non-violent conflict (e.g., rent-seeking activities). Hence, the result suggests that not only the abundance of resources per se but also the lack of strong political and economic institutions (e.g., weak rule of law), which leads to abundant contested resources, are hindrances to good outcomes. Further, an exogenous change making ethnic differences less salient in people’s minds also has effects similar to a rise of the national status.

The results are consistent with the classic thesis on the effects of modernization on social identity, if the national status compared with the ethnic status is high, contested resources are not abundant, institutions are good in quality, or ethnic differences are not salient, otherwise, the results are consistent with the competing thesis, as far as the relatively long-term effect (the effect involving the equilibrium shift) is concerned. Under the former conditions, policies promoting modernization, such as policies stimulating the technological progress of the modern sector and the construction of transportation infrastructure connecting rural areas to urban areas, might be enough for good outcomes. By contrast, under the latter conditions, these policies have negative effects on national identity and conflict and are not very effective for development, and policies raising the national status, improving institutional quality, or making ethnic differences less salient in people’s minds become crucial. Nation-building policies, such as school education or government
propaganda emphasizing common history, culture, and values and the promotion of a national language, may be interpreted as policies raising the national status or making ethnic differences less pronounced and thus are critical under such conditions.

Empirical works suggest the negative effects of natural resources on civil conflict and development and the important effects of institutions on civil conflict, rent-seeking activities, and development. The model reveals a novel mechanism interacting with social identity by which resources and institutions affect conflict and development.

This paper belongs to the literature that uses contest models to examine issues on ethnic conflicts. Aside from Sambanis and Shayo (2013), recent contributions include Esteban and Ray (2008, 2011), Besley and Persson (2010), Sambanis, Skaperdas, and Wohlforth (2015), and Mariani, Mercier, and Verdier (2018). Besley and Persson (2010) examine the interaction between conflict and development as this paper but in connection with capacities of the state to raise revenue and provide market-supporting services. Partly drawing on the model of Sambanis and Shayo (2013), Sambanis, Skaperdas, and Wohlforth (2015) present a model in which the decisions of two countries to go to war or not depend on the expected effect of victory on social identification through a rise in the national status in one of the countries with two opposing social groups.

The paper also belongs to the theoretical literature examining interactions between identity and economic or political behaviors, which includes Fearon and Laitin (2000), Akerlof and Kranton (2000, 2010), Shayo (2009), Benabou and Tirole (2011), Bisin et al. (2011), Gennaioli and Tabellini (2019), and Grossman and Helpman (2020). Akerlof and Kranton (2000) pioneer formally modeling and examining the effects of identity on economic behaviors and Akerlof and Kranton (2010) illustrate how various behaviors can be explained by their framework. By generalizing Akerlof and Kranton (2000), Shayo (2009) constructs the basic framework on which Sambanis and Shayo (2013) and this paper are based and applies it to analyze the political economy of income redistribution. The framework of Shayo (2009) has been applied to various issues. For example, motivated by a recent reversal of trade policies in some western countries seemingly influenced by rises of populism and of ethnic tensions, Grossman and Helpman (2020) construct a political economy model of trade policy with social identification and examine how changes in identification patterns triggered by events such as increased ethnic tensions affect policies.

Finally, the paper is also related to the literature that theoretically examines the modernization of an economy, such as Lewis (1954), Banerjee and Newman (1998), Proto (2007), Vollrath (2009), and Yuki (2007, 2008, 2016). To examine interactions among modernization, conflict, and social identification with a tractable model, this paper models the inefficient sectoral allocation of workers in a simplest manner and considers the modernization induced by exogenous productivity growth. By contrast, these papers model factors leading to the inefficient allocation more explicitly and examine economic mechanisms of modernization more in detail.

5Besides the already mentioned works, recent empirical and experimental studies on identity in economics and political science include Chen and Li (2009), Benjamin, Choi, and Strickland (2010), Eifert, Miguel, and Posner (2010), Clots-Figueras and Masella (2013), Charnysh, Lucas, and Singh (2015), Cohn, Maréchal, and Noll (2015), Alenzuela and Michelson (2016), Benjamin, Choi, and Fisher (2016), and Wimmer (2017), some of which are mentioned later.
The paper is organized as follows. Section 2 presents the model. Section 3 presents and discusses results. In particular, Sections 3.1 and 3.2 examine homogenous and heterogeneous identity equilibria respectively, Section 3.3 analyzes interactions among modernization, identity, conflict, and output, Section 3.4 analyzes the effects of resources on the interactions, and Section 3.5 discusses how results depend on several assumptions of the model. Section 4 concludes. Appendix A presents existence conditions for equilibria, and Appendix B contains proofs.

2 Model

Consider a contest model of conflict in which $n_e (\geq 2)$ ethnic groups contest for exogenous resources. The society is populated by a finite number $L$ of individuals who belong to one of the ethnic groups that are symmetric in every aspect (thus the population size of each group is $L/n_e$). Hence, the model is concerned with an ethnically heterogeneous society without a dominant ethnic group.

**Production:** There are $n_e + 1$ sectors producing private goods: $n_e$ ethnically-segregated traditional sectors (sectors $TJ$, $J = 1, 2, ..., n_e$) and one ethnically-integrated modern sector (sector $M$). The traditional sectors correspond to sectors using traditional or rudimentary technologies in the real economy, such as traditional agriculture, urban informal sector, and household production, and the modern sector corresponds to sectors, such as modern manufacturing and services.\(^6\) The former sectors are more ethnically segregated than the latter (Ezcurra and Rodríguez-Pose, 2017, Beegle et al., 2014):\(^7\) traditional agriculture is operated in largely ethnically homogenous rural communities and typical jobs in the urban informal sector are neighborhood jobs in ethnically segregated communities.

Hence, the production functions of sectors $TJ$ ($J = 1, 2, ..., n_e$) and $M$ are represented as

$$Y_{TJ} = A_T(L_{TJ})^\alpha, \quad \alpha \in (0, 1),$$

$$Y_M = A_M \sum_{J=1}^{n_e} L_{MJ},$$

where $L_{TJ}$ and $A_T$ are the number of workers in sector $TJ$ and the sector’s total factor productivity (TFP) respectively, $L_{MJ}$ is the number of workers of ethnic group $J$ in sector $M$, and $A_M$ is the sector’s TFP. (Each worker supplies a unit of labor inelastically.) Decreasing returns to labor is assumed for sector $TJ$ to capture the fact that labor productivity tends to fall quickly with the amount of labor input in traditional sectors due to limited arable land (traditional agriculture), limited capital available to credit constrained producers (the urban informal sector), or a decreasing

---

\(^6\)The urban informal sector is a part of the urban economy composed of small-scale businesses supplying basic services (for example, small shops and vendors selling commodities and meals) and basic manufacturing goods.

\(^7\)Ezcurra and Rodríguez-Pose (2017) find that ethnic segregation tends to be lower in countries with higher urbanization rates. The review article by Beegle et al. (2014) argue that the majority of informal sector firms are individually-operated or family owned and ethnic and kinship networks play an important role in organizing the sector.
degree of task specialization of each family member (household production).\textsuperscript{8,9}

The wage rate is determined competitively in sector $M$. By contrast, in sector $TJ$, as in Lewis (1954) and many subsequent works modeling traditional sectors, labor income is determined so that the product is equally shared among workers.\textsuperscript{10} Thus, labor incomes in the sectors are

\begin{align*}
y_{TJ} &= A_T(L_{TJ})^{\alpha-1}, \\
y_M &= A_M. \tag{4}
\end{align*}

This setting can generate in a simplest manner the situation facing actual developing countries in which inefficiently many workers exist in traditional sectors and their shift to modern sectors raises aggregate output (Gollin, Lagakos, and Waugh, 2014).\textsuperscript{11}

Individuals of group $J$ are freely mobile between sector $M$ and sector $TJ$, and their sectoral allocation is determined so that their utilities in the two sectors are equated.

**Conflict:** The ethnic groups contest for exogenous resources of value $V$ that are used to produce group-specific club goods, e.g., public services and infrastructures benefiting a particular group.\textsuperscript{12} The amount of resources each group acquires depends on the contributions to the conflict by individuals of each group. In particular, the contested resources are divided among the groups according to the following contest function:

\begin{equation}
\frac{V_J}{V} = \frac{F_J}{F} \text{ if } F > 0 \text{ and } = \frac{1}{n_e} \text{ if } F = 0, \tag{5}
\end{equation}

where $V_J$ is the resources acquired by group $J$ ($J = 1, 2, ..., n_e$), $F_J = \sum_{i \in J} f_i$ is the total contributions or "efforts" by members of the group ($f_i$ is the contribution by individual $i$), and $F = \sum_{J=1}^{n_e} F_J$ is the aggregate "efforts" of the society and is termed the level of conflict. The contested resources represent not only material resources (such as natural resources) but also a part of the governmental budget that can be used to benefit particular ethnic groups.\textsuperscript{13} The model describes a country in which the resource allocation among the groups is determined not by rules but by the consequences of violent or non-violent conflict (such as rent-seeking activities), in which force, mass demonstrations, bribery, or lobbying are employed to influence the outcome.

\textsuperscript{8}This is because the number of tasks performed by each family member increases, as more production activities shift from the market to the household and thus the labor input in household production increases.

\textsuperscript{9}Aside from household production, the assumption of decreasing returns to labor of sector $TJ$ reflects limited amounts of complementary production factors in the real economy. Because such constraint is not severe in the modern sector, constant returns to labor is assumed for sector $M$ for simplicity.

\textsuperscript{10}This assumption reflects the fact that typical production units of traditional sectors are family-run farms/firms or households. Except for the results on the total output of private goods and the aggregate material payoff, the qualitative results below do not depend on this.

\textsuperscript{11}In the real economy, there are other factors causing the inefficient allocation of workers, including inadequate access to quality education required in many modern sector jobs and inadequate access to capital to start a business in the sector. To make the model analytically tractable, these factors are not modeled but would not affect the results.

\textsuperscript{12}The main results would not be affected under the assumption that contested resources yield private goods.

\textsuperscript{13}To be more exact, when the contested resources represent the governmental budget for the club goods, taxation should be modeled. If the government imposes a lump-sum tax of the same amount on all individuals, none of the results are affected. The only change is that the disposable incomes of individuals decrease by the tax payment.
Individual \( i \) contributing \( f_i \) to the conflict incurs a cost of \( c(f_i) \), which, following Esteban and Ray (2011), takes the following form:

\[
c(f_i) = \frac{1}{\theta}(f_i)^\theta, \quad \theta \geq 2.
\]  

(6)

The restriction \( \theta \geq 2 \) is needed to prove some results (\( \theta > 1 \) is enough for most results).

**Utility:** As in Sambanis and Shayo (2013), the utility of an individual depends not only on his or her *material payoff* (the value of private and club goods consumption minus the cost of conflict) positively, but also negatively on the *perceived distance* to a social group the individual chooses to identify with (either his or her ethnic group or the nation) and positively on the *status* of the social group. That is, an individual suffers a large cognitive cost when he or she is very different from others of the group in relevant aspects, while the individual takes pride in being a member of the group when the status of the group is high. These socio-psychological components are important factors affecting social identification and intergroup behaviors, according to influential theories in social psychology (Tajfel and Turner, 1986; Turner et al., 1987) and empirical evidence (Manning and Roy, 2010; Sambanis, Skaperdas, and Wohlforth, 2015).\(^{14}\)

The *material payoff* of individual \( i \) of ethnic group \( J \) \((J = 1, 2, \ldots, n_e)\) when he or she works in sector \( K \) \((K = TJ, M)\) is

\[
\pi_i = y_{K} - \frac{1}{\theta}(f_i)^\theta + \delta V_J,
\]  

(7)

where \( \delta \) measures the strength of preference for the group-specific club good.

*Social groups* are groups from which an individual chooses one group he or she identifies with, which are his or her ethnic group and the nation \( N \).

Individual \( i \) who is characterized by two types of attributes perceives how close to or far from a social group by the distance between his or her attributes and the average attributes of the group. The attributes are whether the individual belongs to (a) particular ethnic groups or not, and (b) particular traditional sectors or not.\(^{15}\)

\(^{14}\) Evidence suggests that perceived distance and status affect social identity. For example, Manning and Roy (2010) find for the UK that nonwhites, whose perceived distance to the "average person" in the nation seems to be greater, are less likely to think of themselves as British than whites. Further, they find that immigrants from poorer and less democratic (i.e., lower status) countries assimilate faster into a British identity. Sambanis, Skaperdas, and Wohlforth (2015) present episodes suggesting that interstate wars affect social identity through the national status, including increased identification with the state after the victory in the WWII in the USSR and the intensification of a common identity among southern Slavs, including Croats and Slovenes, after Serbian victories against the Ottoman Empire and Bulgaria in the Balkan Wars in 1912–13.

\(^{15}\) The reason why sectoral affiliation is a component of the perceived distance is that the type of job (modern sector job or traditional sector job) and the region of residence (urban or rural region) of an individual are considered as important factors affecting the person’s social identity. In the second round of the Afrobarometer, a multicounty African public attitude survey project, conducted in 2002-03 (https://www.afrobarometer.org/online-data-analysis/analyse-online), there is a question asking respondents to which group they feel belong foremost, besides the nation. (This is the only round asking such question.) The highest proportion of respondents choose occupation (27.2%) rather than ethnic group (22.6%). YouGov, albeit a British survey, has a similar question that, unlike the Afrobarometer, allows for multiple answers and the two most common answers are occupation (44%) and the area of residence (43%) (Shayo, 2009).
\[ q_i^J = 1 \text{ if } i \in J; \quad q_i^J = 0 \text{ otherwise, for } J = 1, 2, \ldots, n_e, \quad (8) \]
\[ q_i^{TJ} = 1 \text{ if } i \in TJ; \quad q_i^{TJ} = 0 \text{ otherwise, for } J = 1, 2, \ldots, n_e. \quad (9) \]

For example, when the person belongs to ethnic group 2 and works in sector \( M \), \( q_i^2 = 1, q_i^J = 0 \) for \( J \neq 2 \), and \( q_i^{TJ} = 0 \) for any \( J \). Note that the ethnic attributes are fixed, while the sectoral attributes are determined endogenously by sectoral choices of workers, which as shown later, could affect their identity choices.

The perceived distance between individual \( i \) and social group \( G (G = J, N) \), on which his or her utility negatively depends, is represented by\(^{16}\)

\[ d_{iG}^2 = \omega_e \sum_{J=1}^{n_e} (q_i^J - q_G^J)^2 + \omega_s \sum_{J=1}^{n_e} (q_i^{TJ} - q_G^{TJ})^2, \quad (10) \]

where \( q_G^J \) and \( q_G^{TJ} \) are respectively the average values of the ethnic and sectoral attributes of the group; \( \omega_e \) and \( \omega_s \) are weights on the respective types of attributes. The first term, which is henceforth called the ethnic distance, is the psychological cost the individual suffers due to the difference in ethnicity between oneself and the "average member" of the group. The second term, which is named the sectoral distance, is the psychological cost due to the difference in sectoral affiliation, which represents the type of job or the region of residence in the real society (see footnote 15).

Following Sambanis and Shayo (2013), the weight on the ethnic attributes \( \omega_e \) is assumed to increase with the level of ethnic conflict \( F \):

\[ \omega_e = \eta_0 + \eta_1 F, \quad \eta_0 \geq 0, \quad \eta_1 > 0, \quad (11) \]

The specification implies that when ethnic conflict becomes more intense, people care about the ethnic attributes more in measuring distances from social groups. This is consistent with empirical evidence: A case analysis of the civil war in Yugoslavia in the 1990s by Sambanis and Shayo (2013) cites evidence showing that the share of people identifying themselves as “Yugoslavs” dropped greatly after the intensification of the conflict despite episodes suggesting the lack of strong ethnic identities before the war; Rohner, Thoenig, and Zilibotti (2013), using individual, county-level and district-level data from Uganda, find that the proportion of individuals identifying more with their ethnic group than with the nation is higher in counties of a higher intensity of armed conflicts, after controlling for individual, ethnic, and spatial characteristics and employing instrumental variable estimation; Eifert, Miguel, and Posner (2010), based on 22 public opinion surveys in 10 African countries, find that being close to a competitive presidential election is positively associated with ethnic identification.

The utility of an individual also depends positively on the status of the social group \( G (G = J, N) \) he or she identifies with, which is exogenous and denoted by \( S_G \). Ethnic groups are symmetric

\(^{16}\)The concept of perceived distance is developed in cognitive psychology in studying how a person categorizes information that comes in to his or her (stimuli) (Nosofsky, 1986). Turner et al. (1987) apply the concept to the categorization by a person of people, including oneself, into social groups, in constructing an influential social psychological theory, self-categorization theory. The theory tries to explain the psychological basis of social identification.
and thus the level of their status is the same and is denoted by

\[ S_J = S_E \] for every \( J \).

The level of the national status \( S_N \) would be high when the people of a nation believe that they share glorious history, rich culture, or a "right" sense of values, or when the nation has great performance in international sports, because the people would feel proud of belonging to such nation. By contrast, the level of the ethnic status \( S_E \) would be high when ethnic groups are distinctive in these dimensions.\(^{17,18}\)

From these settings, the utility of individual \( i \) who identifies with social group \( G \) is given by

\[ u_{iG} = \pi_i - \beta d_{iG}^2 + \gamma S_G, \quad \beta, \gamma > 0. \] (13)

The utility function implies that given that an individual identifies with a particular social group, his or her utility increases as the perceived distance to the group falls; thus, the individual has an incentive to choose actions lowering the distance. For example, since the sectoral distance is a component of the perceived distance, others things equal, the individual has an incentive to choose the same sector as the "average person" of the group.

However, social identification of an individual, that is, the group he or she identifies with, is not fixed. The individual "chooses" a group (his or her ethnic group or the nation) that brings him or her higher utility because of a higher material payoff, a shorter perceived distance, or a higher status.\(^{19}\) His or her social identity might change if the exogenous variables affecting the utility directly or indirectly through choices by others alter. For example, as the level of conflict rises, people place a greater weight on ethnicity in the perceived distance, which could change their identities. The exact timing of their decisions is as follows.

**Timing:** Individuals play a two-stage game to maximize their utility. First, they decide in which sector to work (sector \( T_J \) or sector \( M \) for ethnic group \( J \)), which determines the labor income in the traditional sectors \((y_{TJ})\) and sectoral and aggregate output \((Y_{TJ}, Y_M, \text{and } Y \equiv Y_{TJ} + Y_M)\). Then, that is, after \( L_{TJ} \) and \( L_{MJ} \) are determined, they simultaneously choose a group to identify with and their contribution to conflict \( f_i \), which determines the level of conflict \( F \), the allocation of contested resources among the groups, and individual utilities.\(^{20}\) The solution concept applied is

\(^{17}\)Similar to works such as Grossman and Helpman (2020), status is an absolute measure. By contrast, in Shayo (2009) and Sambanis and Shayo (2013), status is a relative measure and is defined as the difference from the reference group. Main results remain unchanged under the alternative specification.

\(^{18}\)To make the model manageable, unlike Sambanis and Shayo (2013) and Grossman and Helpman (2020), the status does not depend on the group’s total material payoffs (the sum of \( \pi_i \)). The results would not be affected by considering the economic status, as long as its importance in the utility is not very large.

\(^{19}\)By assumption, the individual does not identify with both the ethnic group and the nation. In contrast, in the model of Grossman and Helpman (2019), an individual always identifies with his or her socioeconomic class and also identifies with the nation if the additional identity increases the utility, where the utility depends on the sum of the perceived distance to and the status of each group the individual identifies with. Such specification is not adopted in this paper, because the perceived distance term becomes complicated and analysis of the model becomes difficult.

\(^{20}\)The timing of events reflects the fact that the sectoral choice made earlier in life largely determines the sector to work in for most of life (because, in the real economy, the sectors tend to require different levels of education and different types of skills and tend to be located in different places), while social identity is more likely to change over
subgame perfect Nash equilibrium; thus, the two-stage game can be solved by backward induction.\footnote{Sambanis and Shayo (2013) apply the concept of the social identity equilibrium to their one-shot game. The equilibrium is similar to the standard Nash equilibrium, but the condition on the choice of identities is weaker. In this paper, the concept of the subgame perfect Nash equilibrium is used, because it is more familiar and is easier to apply. Shayo (2009) also employs the standard Nash equilibrium to solve a one-shot game of social identity.}

## 3 Results

As mentioned earlier, ethnic groups are assumed to be symmetric in every aspect. Hence, the paper focuses on the equilibria in which choices of all groups are symmetric.\footnote{There also exist subgame perfect Nash equilibria in which different ethnic groups make different choices, which are generally very difficult to analyze.} These equilibria can be classified into two types: equilibria in which individuals of the same ethnic group share the same identity and those in which they do not. For ease of exposition, homogenous identity equilibria are analyzed first (Section 3.1); then, heterogenous identity equilibria are analyzed and compared with homogenous identity equilibria (Section 3.2). These sections compare different equilibria for given parameters and exogenous variables, but the set of realized equilibria changes with values of the exogenous variables. Taking this into account, Section 3.3 analyzes the focus of the paper: the interactions among modernization, identity, conflict, and output. Section 3.4 examines how the abundance of contested resources affects the interactions.

To simplify the analysis, the following assumption, which is a sufficient condition for \( f_i > 0 \) and thus \( F > 0 \) to hold in all equilibria, is imposed.

\begin{equation}
V_L > (\beta \eta_1)^{\frac{\theta}{\gamma-1}} \left( \frac{n_e}{n_e} \right)^{\frac{1}{\gamma-1}}.
\end{equation}

### 3.1 Homogenous identity equilibria

There exist two homogenous identity equilibria: the equilibrium in which all individuals identify with their ethnic group (henceforth \textit{equilibrium (e)}) and the one in which all individuals identify with the nation (henceforth \textit{equilibrium (n)}).

#### 3.1.1 Equilibrium (e)

Consider the second stage of the game in which the sectoral allocation of workers (\( L_{TJ} \) and \( L_{MJ} \)) is given. When individual \( i \) of ethnic group \( J \) (\( J = 1, 2, ..., n_e \)) in sector \( M \) identifies with his or her ethnic group, the individual chooses the contribution to conflict \( f_i \) to maximize the following utility (note \( q_{iJ} = q_{J} = 1, q_{i}^k = 0 \) for \( k \neq J, q_{J}^{TJ} = \frac{L_{TJ}}{L/n_e}, \) and \( q_{J}^{k} = 0 \) for \( k \neq J, TJ \)):

\begin{equation}
A_M - \frac{1}{\theta} (f_i)^{\theta} + \delta \frac{F_j}{F} V - \beta \omega_s \left( \frac{L_{TJ}}{L/n_e} \right)^2 + \gamma S_E.
\end{equation}
\[ f_i = f_{i,e} = \left( \frac{\delta F_i J}{F^2} V \right)^{\frac{1}{\gamma-1}}, \text{ where } F_{-j} = F - F_j. \] (16)

When the individual is in sector \( TJ \) instead, he or she chooses \( f_i \) to maximize \( q_i^{TJ} = 1 \)
\[ A_T(L_{TJ})^{\alpha-1} - \frac{1}{\theta}(f_i)^\theta + \delta \frac{F_i}{F} V - \beta \omega_s \left( 1 - \frac{L_{TJ}}{L/n_e} \right)^2 + \gamma S_E. \] (17)
The solution for \( f_i \) is given by (16) as in the previous case.

Since all individuals identify with their ethnic group and the ethnic groups are symmetric, by substituting \( F_{-j} = \frac{n_e}{n_e} F \) and \( f_i = F/L \) into (16), the equilibrium level of conflict \( F_e^* \) is obtained:
\[ F_e^* = \left( \frac{n_e - 1}{n_e} \right)^{\frac{1}{\gamma}} L \text{ from } F_e^* = \left( \frac{n_e - 1}{n_e} \right)^{\frac{1}{\gamma-1}} L. \] (18)

In the first stage, individuals choose sectors by taking into account the effects of their choices on the second stage. Assume that the following condition holds so that \( L_{TJ} = \frac{L}{n_e} \) (all individuals choose sector \( TJ \)) does not hold in equilibrium:

Assumption 2: \[ A_T \left( \frac{L}{n_e} \right)^{\alpha-1} + \beta \omega_s < A_M. \] (19)

Then, the sectoral allocation of workers is determined so that choosing either sector is indifferent. From (15) and (17), the indifference condition is
\[ A_T(L_{TJ})^{\alpha-1} - \beta \omega_s \left( 1 - 2n_e \frac{L_{TJ}}{L} \right) = A_M, \] (20)
which gives the unique \((L_{TJ})^*_e \in (0, \frac{L}{n_e})\) that decreases with \( A_M \) and increases with \( A_T \).²³

### 3.1.2 Equilibrium (n)

Consider the second stage of the game in which the sectoral allocation of workers is given. When individual \( i \) of ethnic group \( J \) in sector \( M \) identifies with the nation, he or she chooses \( f_i \) to maximize the following utility \( (\omega_e = \eta_0 + \eta_i F, q_{i,J}^M = \frac{n_e}{n_e}, n_{TK} = \frac{L_{TK}}{L} \text{ for any } TK): \)
\[ A_M - \frac{1}{\theta}(f_i)^\theta + \delta \frac{F_i}{F} V - \beta \left( \eta_0 + \eta_i F \right) \left( \frac{n_e - 1}{n_e} + \omega_s \left( \frac{L_{TJ}}{L} \right)^2 \right) \] + \gamma S_N. \] (21)

From the first-order condition \( (f_i > 0 \text{ from } (14)), \)
\[ f_i = f_{i,n} = \left( \frac{\delta F_i J}{F^2} V - \beta \eta_i \frac{n_e - 1}{n_e} \right)^{\frac{1}{\gamma-1}}, \text{ where } F_{-j} = F - F_j. \] (22)

When the individual is in sector \( TJ \) instead, he or she chooses \( f_i \) to maximize

²³ The first derivative with respect to \( L_{TJ} \) of the LHS of (20) is \(-\frac{1}{L_{TJ}} \frac{\partial A_T(L_{TJ})^{\alpha-2} + \beta \omega_s \frac{2n_e}{L}}{L_{TJ}} \), which equals \(-\infty \) at \( L_{TJ} = 0 \) and equals \( 0 \) at \( L_{TJ} = \frac{\left( 1 - \alpha \right) A_T (L_{TJ})^{\alpha-3}}{\beta \omega_s L} \), and the second derivative equals \( 2 - \alpha \left( 1 - \alpha \right) A_T (L_{TJ})^{\alpha-3} > 0 \). Thus, from (19) and the fact that the LHS of (20) at \( L_{TJ} = 0 \) equals \( +\infty \), there exists a unique \( L_{TJ}^* \in (0, \frac{L}{n_e}) \) satisfying (20). The relations of \((L_{TJ})^*_e \) with \( A_M \) and \( A_T \) are straightforward from the shape of the LHS of (20).
\[ A_T(L_{TJ})^{\alpha-1} - \frac{1}{\theta} (f_i)^\theta + \delta \frac{F_j}{F_j} V - \beta \left( \eta_0 + \eta_1 F \right) \frac{n_e - 1}{n_e} \omega_s \left( \frac{1 - L_{TJ}}{L} \right)^2 + \sum_{K \neq J} \left( \frac{L_{TK}}{L} \right)^2 \right) + \gamma S_N, \] (23)

whose solution is given by (22).

Since all individuals identify with the nation and the groups are symmetric, by plugging \( F_j = n_e F \) and \( f_i = F/L \) into (22), the equilibrium level of conflict \( F_n^* \) is obtained as a solution for

\[ F_n^* = \left[ \frac{n_e - 1}{n_e} \left( \delta \frac{V}{F_n^*} - \beta \eta_1 \right) \right]^{\frac{1}{\theta - 1}} L. \] (24)

In the first stage, from (21) and (23), the indifference condition for sectoral choices equals

\[ A_T(L_{TJ})^{\alpha-1} - \beta \omega_s \left( 1 - 2 \frac{L_{TJ}}{L} \right) = A_M, \] (25)

which gives the unique \((L_{TJ})_n^* \in (0, (L_{TJ})_e^*)\) that decreases with \( A_M \) and increases with \( A_T \).\(^{24}\)

### 3.1.3 Analysis

Based on Proposition A1 in Appendix A, Figure 1 illustrates combinations of \( L_{TJ} \) and \( \Delta S \equiv S_N - S_E \) such that homogenous identity equilibria exist. In the figure, equilibrium (n) exists when \((L_{TJ})_n^* \) and \( \Delta S \) are located in the region above the upward-sloping curve, and equilibrium (e) exists when \((L_{TJ})_e^* \) and \( \Delta S \) are located in the region on or below the downward-sloping curve. Given \( L_{TJ} \), equilibrium (n) (equilibrium (e)) exists when \( \Delta S \) is sufficiently high (low). This is because the utility under the national identity is large (small) relative to the utility under the ethnic identity when the national status compared with the ethnic status is high (low). Given \( \Delta S \), at least one of

\(^{24}\)From the comparison of the LHS of (25) with that of (20) and the discussion in footnote 23, it is clear that, when (19) is assumed, the unique solution \((L_{TJ})_n^* \in (0, (L_{TJ})_e^*)\) that decreases with \( A_M \) and increases with \( A_T \) exists.
the equilibria exists when $LTJ$ is small enough. Homogenous identity equilibria do not exist when the number of traditional sector workers is sufficiently large, because modern sector workers do not identify with their ethnic group due to a large sectoral distance to the "average co-ethnic" compared to their distance to the "average national", while traditional sector workers do not identify with the nation due to a large sectoral distance to the "average national" compared to their distance to the "average co-ethnic". Note that both equilibria exist in the region with slant lines.\footnote{The reason for multiple equilibria is explained when the mechanism of the main result is explained in Section 3.3.1.}

How do the two equilibria differ in terms of conflict, the sectoral distribution of workers, and output? The next proposition compares them for given parameters and exogenous variables.

**Proposition 1** Suppose that the two homogenous identity equilibria exist for some given parameters and exogenous variables. Then, the following holds.

(i) The level of conflict is lower in equilibrium (n), i.e., $F_n^* < F_e^*$.  
(ii) $LTJ$ and thus the proportion of workers in the traditional sectors are lower in equilibrium (n), i.e., $(LTJ)_n^* < (LTJ)_e^*$.  
(iii) The output of private goods $Y$ is higher in equilibrium (n), if $\alpha$ (the parameter of the traditional sector production function) or $\beta$ (the importance of the perceived distance in the utility) is not very high. The output is higher in equilibrium (e) if $\alpha$ or $\beta$ is high.  
(iv) Aggregate material payoff is higher in equilibrium (n), unless $\alpha$ or $\beta$ is very high.

People contribute less to conflict and the level of conflict $F$ is lower when they identify with the nation, because, in choosing $f_i$, they take into account the undesirable effect of conflict on the ethnic distance to the "average national": a higher $F$ raises the weight on ethnicity, $\omega_e$, thereby highlighting the differences among citizens and raising the distance. This result is shown in Sambanis and Shayo (2013) and consistent with above-mentioned evidence (Eifert, Miguel, and Posner, 2010; Rohner, Thoenig, and Zilibotti, 2013).

What is new is the effect on the sectoral distribution of individuals and total output. For given parameters and exogenous variables, $LTJ$ and thus the proportion of workers in the traditional sectors are lower in equilibrium (n). The return from choosing the traditional sector of their ethnic group is lower under national identity, because the sectoral distance to the "average national" rises by choosing the ethnically segregated sector over the integrated modern sector, whereas, under ethnic identity, the sectoral distance to the "average co-ethnic" falls (if $LTJ > \frac{L}{2n_e}$, i.e., the majority is in the traditional sector) or rises less (if $LTJ < \frac{L}{2n_e}$) by choosing the traditional sector.

The sectoral allocation of workers is generally inefficient: i.e., it does not maximize the total output of private goods, because labor income is greater than the marginal labor productivity in the traditional sectors (note decreasing returns to labor, $\alpha < 1$) and the sectoral distance term in the utility distorts sectoral choices by inducing workers to choose the same sector as the "average fellow" of their identity group. The former leads to too many traditional sector workers, while the latter leads to too few traditional sector workers in equilibrium (n) and leads to too many (few) workers in the traditional sectors in equilibrium (e) when $LTJ > (<) \frac{L}{2n_e}$.
If $\alpha$ is not very high, the first effect dominates and thus $L_{TJ}$ is higher than the efficient level. Total output is higher in equilibrium (n) because $L_{TJ}$ is smaller and thus closer to the efficient level than in equilibrium (e). The condition would be more relevant to developing nations, since a small $\alpha$ implies strong decreasing returns in the traditional sectors.\textsuperscript{26} The same result holds for any $\alpha$, if the importance of the perceived distance in the utility, $\beta$, is not very high and thus the effect of social identity on sectoral misallocation (the second effect) does not exceed the economic effect (the first effect), which is likely to be true.\textsuperscript{27} Finally, aggregate material payoff (the value of private and club good consumption net of the cost of conflict) is higher in equilibrium (n) because of the lower cost of conflict, unless $\alpha$ or $\beta$ is very high.\textsuperscript{28}

To summarize, national identity is associated with not only a lower level of conflict but also higher shares of modern sector workers and production and under plausible conditions, it is also associated with higher levels of the total output of private goods and of aggregate material payoff. Note that the result for the output and material payoff is obtained even though the model does not assume the plausible negative effect of conflict on modern sector productivity (relative to traditional sector productivity). The result would be strengthened if such effect is considered.\textsuperscript{29}

### 3.2 Heterogeneous identity equilibria

Now, equilibria in which individuals of the same ethnic group have different identities are examined. There exist three heterogeneous identity equilibria: the equilibrium in which sector $TJ$ workers identify with their ethnic group and sector $M$ workers identify with the nation (henceforth equilibrium (d); "d" is for divided identities); the one in which sector $M$ workers are divided over identities and sector $TJ$ workers identify with their ethnic group (henceforth equilibrium (Md)); and the one in which sector $TJ$ workers are divided over identities and sector $M$ workers identify with the nation (henceforth equilibrium (Td)). Readers who are not interested in derivations of equations for these equilibria may directly go to the analysis of Section 3.2.4.

#### 3.2.1 Equilibrium (d)

In the second stage in which the sectoral allocation of workers is given, sector $TJ$ workers with an ethnic identity choose $f_i$ to maximize (17), and the solution is given by (16), while sector $M$ workers with a national identity choose $f_i$ to maximize (21), and the solution is given by (22).

Because the ethnic groups are symmetric, by substituting $F_{J} = n_e - \frac{1}{n_e} F$ into (16) and (22) and

\textsuperscript{26}Remember that the decreasing returns to labor intends to capture the fact that labor productivity tends to fall with the amount of labor input in the sectors due to limited arable land (traditional agriculture), limited capital available to credit constrained producers (the urban informal sector), or a decreasing degree of task specialization of each family member (household production).

\textsuperscript{27}By contrast, $Y$ is lower in equilibrium (n), if $\alpha$ or $\beta$ is high enough that the second effect dominates and thus $L_{TJ}$ is lower than the efficient level.

\textsuperscript{28}When $\alpha$ or $\beta$ is very high, $Y$ is much lower in equilibrium (n) (footnote 27). Then, despite the lower cost of conflict, aggregate material payoff is lower in equilibrium (n).

\textsuperscript{29}The easiest way to include this effect is to assume that $A_M(F), A_M'(F) < 0$, and individuals do not consider effects of their actions on $A_M(F)$ in making decisions. Then, only the indifference conditions for sectoral choices change.
plugging them into $F = f_{i,e}n_eL_{TJ} + f_{i,n}(L - n_eL_{TJ})$, the level of conflict $F$ for given $L_{TJ}$ is obtained:

$$F = \left(\frac{n_e - 1}{n_e}\right)^{\frac{1}{\theta - 1}} \left[ (\delta V F)^{\frac{1}{\theta - 1}} n_eL_{TJ} + (\delta V F - \beta \eta_1)^{\frac{1}{\theta - 1}} (L - n_eL_{TJ})\right], \quad (26)$$

which increases with $L_{TJ}$ and is denoted by $F_d(L_{TJ})$ ($d$ is for "divided identities").

In the first stage, from (16), (17), (21), (22), and (26), the indifference condition for sectoral choices equals ($\Delta S \equiv S_N - S_E$)

$$A_T(L_{TJ})^{\alpha - 1} - \frac{1}{\theta} \left(\delta \frac{(n_e - 1)}{n_e} \frac{V F}{F_d(L_{TJ})}\right)^{\frac{\theta}{\theta - 1}} - \beta \omega_s \left(1 - \frac{L_{TJ}}{L/n_e}\right)^2 + \gamma S_E$$

$$= A_M - \frac{1}{\theta} \left[\frac{n_e - 1}{n_e} \left(\delta \frac{V F}{F_d(L_{TJ})} - \beta \eta_1\right)\right]^{\frac{\theta}{\theta - 1}} - \beta \left[\eta_0 + \eta_1 F_d(L_{TJ}) \frac{n_e - 1}{n_e} + \omega_s n_e \left(\frac{L_{TJ}}{L}\right)^2\right] + \gamma S_N$$

$$\Leftrightarrow A_T(L_{TJ})^{\alpha - 1} + \beta \left[\eta_0 + \eta_1 F_d(L_{TJ}) \frac{n_e - 1}{n_e} + \omega_s \left(\frac{L_{TJ}}{L}\right)^2\right] - \frac{1}{\theta} \left[\frac{n_e - 1}{n_e}\right]^{\frac{\theta}{\theta - 1}} \left(\delta \frac{V F}{F_d(L_{TJ})} - \beta \eta_1\right)^{\frac{\theta}{\theta - 1}} - \gamma \Delta S = A_M, \quad (27)$$

which gives solution $(L_{TJ})^*_d \in (0, \frac{L}{n_e})$. Appendix B proves its uniqueness when $\theta = 2$.\(^{30}\) The equilibrium level of conflict, $F^*_d$, is obtained from the substitution of $(L_{TJ})^*_d$ into (26).

### 3.2.2 Equilibrium (Md)

In the second stage, workers in sector $TJ$ identifying with their ethnic group choose $f_i$ to maximize (17), and the solution is given by (16); workers in sector $M$ are indifferent between identifying with the nation, in which case $f_i$ is chosen to maximize (21) and the solution is given by (22), and identifying with their ethnic group, in which case $f_i$ is chosen to maximize (15) and the solution is given by (16).

Thus, the indifference condition for *identity* choices when ethnic groups are symmetric is

$$A_M - \frac{1}{\theta} (f_{i,e})^\theta + \delta \frac{F_d}{F} V - \beta \omega_s \left(\frac{L_{TJ}}{L/n_e}\right)^2 + \gamma S_E$$

$$= A_M - \frac{1}{\theta} (f_{i,n})^\theta + \delta \frac{F_d}{F} V - \beta \left[\eta_0 + \eta_1 F \frac{n_e - 1}{n_e} + \omega_s n_e \left(\frac{L_{TJ}}{L}\right)^2\right] + \gamma S_N$$

$$\Leftrightarrow \beta \left[\eta_0 + \eta_1 F \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \left(\frac{L_{TJ}}{L}\right)^2\right] - \frac{1}{\theta} \left[\frac{n_e - 1}{n_e}\right]^{\frac{\theta}{\theta - 1}} \left[\delta \frac{V F}{F_d(L_{TJ})} - \beta \eta_1\right]^{\frac{\theta}{\theta - 1}} = \gamma \Delta S. \quad (29)$$

$F$ in the above equation satisfies

$$F = f_{i,n} P_{M,n} (L - n_eL_{TJ}) + f_{i,e} \left[n_eL_{TJ} + (1 - P_{M,n})(L - n_eL_{TJ})\right]$$

$$= \left(\frac{n_e - 1}{n_e}\right)^{\frac{1}{\theta - 1}} \left[\delta \frac{V f_{i,n}}{F_d(L_{TJ})} P_{M,n} (L - n_eL_{TJ}) + \left(\frac{V f_{i,n}}{F_d(L_{TJ})}\right)^{\frac{1}{\theta - 1}} \left[n_eL_{TJ} + (1 - P_{M,n})(L - n_eL_{TJ})\right]\right], \quad (30)$$

where $P_{M,n}$ is the proportion of sector $M$ workers identifying with the nation. Since the LHS of (29) decreases with $L_{TJ}$ and increases with $F$, $F$ satisfying (29) increases with $L_{TJ}$.

---

\(^{30}\)When $\theta > 2$, the uniqueness of $(L_{TJ})^*_d$ cannot be proved, but whether $(L_{TJ})^*_d$ is unique or not does not affect the results below.
The indifference condition for sectoral choices in the first stage is given by (20) and is thus the same as equilibrium (e) from (15) and (17). Thus, the equilibrium level of \( L_{TJ} \), \( (L_{TJ})_{Md}^{*} \), equals \( (L_{TJ})_{d}^{*} \), and the equilibrium level of conflict \( F_{Md}^{*} \) is obtained by substituting \( (L_{TJ})_{e}^{*} \) into (29) and solving it for \( F \).

3.2.3 Equilibrium (Td)

In the second stage, workers in sector \( M \) identifying with the nation choose \( f_{i} \) to maximize (21), and the solution is (22); workers in sector \( TJ \) are indifferent between identifying with the nation, in which case \( f_{i} \) is chosen to maximize (23) and the solution is given by (22), and identifying with their ethnic group, in which case \( f_{i} \) is chosen to maximize (17) and the solution is given by (16).

Thus, the indifference condition for identity choices when ethnic groups are symmetric is

\[
A_T(L_{TJ})^{\alpha-1} - \frac{1}{\beta}(f_{i,e})^{\theta} + \delta \frac{P_{TJ}}{P_{L}} V - \beta \omega_s \left( 1 - \frac{L_{TJ}}{n_e} \right)^{1/2} + \gamma S_E
\]

\[
= A_T(L_{TJ})^{\alpha-1} - \frac{1}{\beta}(f_{i,n})^{\theta} + \delta \frac{P_{TJ}}{P_{L}} V - \beta \left\{ \omega_s \frac{n_e}{n_e} + \omega \left[ (1 - L_{TJ})^{1/2} + (n_e - 1)(L_{TJ})^{1/2} \right] \right\} + \gamma S_N
\]

\[
\Leftrightarrow \beta \left[ \eta_0 + \eta_1 F \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{P_{L}} \left( 2 - n_e \frac{L_{TJ}}{P_{L}} \right) - \frac{1}{\beta} \left( n_e - 1 \right) \right] \right\]^{\frac{\theta}{\theta-1}} \left[ \delta \frac{P_{TJ}}{P_L} \right]^{\frac{\theta}{\theta-1}} - (\delta \frac{P_{TJ}}{P_L} - \beta \eta_1)^{\frac{\theta}{\theta-1}} = \gamma \Delta S. \tag{32}
\]

\( F \) in the above equation satisfies

\[
F = f_{i,n} \left[ P_{TJ,n} n_e L_{TJ} + (L - n_e L_{TJ}) \right] + f_{i,e} \left( 1 - P_{TJ,n} \right) n_e L_{TJ}
\]

\[
= \left( \frac{n_e - 1}{n_e} \right) \left[ \left( \frac{P_{TJ,n}}{P_{L}} \right) \left( L_{TJ} + (L - n_e L_{TJ}) \right) + \left( \frac{P_{TJ}}{P_{L}} \right) \left( 1 - P_{TJ,n} \right) n_e L_{TJ} \right], \tag{33}
\]

where \( P_{TJ,n} \) is the proportion of sector \( TJ \) workers identifying with the nation. \( F \) satisfying (32) decreases with \( L_{TJ} \) because the LHS of (32) increases with \( L_{TJ} \) and \( F \). After the negative dependence of \( F \) on \( L_{TJ} \) is taken into account, \( P_{TJ,n} \) increases with \( L_{TJ} \) from (33).

The indifference condition for sectoral choices in the first stage is given by (25) and thus is the same as equilibrium (n) from (21) and (23). The equilibrium level of \( L_{TJ} \), \( (L_{TJ})_{Md}^{*} \), equals \( (L_{TJ})_{d}^{*} \), and the equilibrium level of conflict \( F_{Md}^{*} \) is obtained by substituting \( (L_{TJ})_{e}^{*} \) into (22) and solving it for \( F \).

3.2.4 Analysis

Based on Proposition A2 in Appendix A, Figure 2 illustrates combinations of \( L_{TJ} \) and \( \Delta S \) such that heterogenous identity equilibria and homogenous identity equilibria exist.\(^{31}\) Equilibrium (d) exists when \( (L_{TJ})_{d}^{*} \) and \( \Delta S \) are located in the region with triple-dashed double-dotted lines; equilibrium (Md) exists when \( (L_{TJ})_{Md}^{*} \) and \( \Delta S \) are located in the region with negatively-sloped lines; and equilibrium (Td) exists when \( (L_{TJ})_{Td}^{*} \) and \( \Delta S \) are located in the region with positively-sloped lines.

\(^{31}\)To be precise, this figure illustrates combinations of \( L_{TJ} \) and \( \Delta S \) when \( \omega_s \) (the weight on sectoral attributes in the perceived distance) is relatively high and \( \eta_1 \) (the strength of the effect of \( F \) on the weight on ethnic attributes) is relatively low. Figure 5 in Appendix A illustrates combinations of \( L_{TJ} \) and \( \Delta S \) when \( \omega_s \) is relatively low and \( \eta_1 \) is relatively high. Basic features of the two figures are similar.
As in Figure 1, equilibrium (n) (equilibrium (e)) exists in the region above the upward-sloping solid curve (on or below the downward-sloping solid curve).

Given $L_TJ$, the heterogeneous identity equilibria exist when $\Delta S$ is neither very high nor very low: when the national status relative to the ethnic status is very high (low), everyone identifies with the nation (their ethnic group). Given $\Delta S$, equilibrium (d) exists when $L_TJ$ is sufficiently large: when many workers are located in traditional sectors, traditional sector workers identify with their ethnic group mainly because the sectoral distance to the "average co-ethnic" is small, while modern sector workers identify with the nation mainly because the sectoral distance to the "average co-ethnic" is large. Note that there are several regions in which multiple equilibria exist.\(^{32}\) Equilibria (Md) and (Td) exist only in the region where equilibrium (d) exists, and all equilibria exist in the small triangular region on the left side of the figure.

In all the heterogeneous identity equilibria, modern sector workers are more (less) likely to identify with the nation (their ethnic group) than traditional sector workers: when some workers in the traditional sectors identify with the nation, so do all workers in the modern sector (and when some modern sector workers identify with their ethnic group, so do all traditional sector workers).\(^{33}\) Roughly speaking, this is because the modern sector is ethnically integrated (and the traditional sectors are ethnically segregated) and thus the sectoral distance to the "average national" of modern sector workers is smaller than their sectoral distance to the "average co-ethnic", and the opposite.

\(^{32}\)The reason for multiple equilibria is explained when the mechanism of the main result is explained in Section 3.3.1.

\(^{33}\)The non-existence of equilibria in which modern sector workers are less likely to identify with the nation than traditional sector workers is formally shown in the proof of Proposition A2 of Appendix A.
holds for traditional sector workers. The result is consistent with Robinson (2014), who, using individual-level survey data of sixteen African countries, finds that being employed in the modern sector is significantly and robustly associated with identifying more with the nation than with their ethnic group, after controlling for education, urban residence, gender, and group-level and country-level variables.

Proposition 1 shows for homogenous identity equilibria that social identity affects conflict, the sectoral distribution of workers, and output. The next proposition compares all equilibria in terms of these variables for given parameters and exogenous variables.

**Proposition 2** Suppose that multiple equilibria exist for some given parameters and exogenous variables. Then, the following holds when these equilibria are compared.

(i) The level of conflict is lower when the proportion of individuals identifying with the nation is higher, i.e., \( F_n^* < F_{Td}^* < F_{Md}^* < F_e^* \).

(ii) \((L_{TJ})_{Td}^* = (L_{TJ})_n^* < (L_{TJ})_d^* < (L_{TJ})_{Md}^* = (L_{TJ})_e^* \).

(iii) \( Y_{Td}^* = Y_n^* > Y_d^* = Y_{Md}^* = Y_e^* \) if \( \alpha \) or \( \beta \) is not very high. The relation is opposite if \( \alpha \) or \( \beta \) is sufficiently high.

(iv) Aggregate material payoff is higher when the proportion of individuals identifying with the nation is higher, unless \( \alpha \) or \( \beta \) is very high.

The level of conflict is lower when the proportion of individuals identifying with the nation is higher. Among the heterogenous identity equilibria, equilibrium (Td) in which all sector M workers identify with the nation and sector TJ workers are divided over identities has the lowest conflict level, and equilibrium (Md) in which sector M workers are divided over identities and sector TJ workers identify with their ethnic group has the highest conflict level. Among all the equilibria, equilibrium (n) and equilibrium (e) have lowest and highest conflict levels respectively.

Roughly speaking, the reason is that those identifying with the nation contribute less to conflict, because, in choosing \( f_i \), they take into account that a higher \( F \) raises the ethnic distance to the "average national" by highlighting differences among ethnic groups.

The above explanation presumes that, among the heterogenous identity equilibria, the proportion of individuals identifying with the nation is highest in equilibrium (Td) and lowest in equilibrium (Md). The result on the fraction of workers in the traditional sectors, \((L_{TJ})_{Td}^* = (L_{TJ})_n^* < (L_{TJ})_d^* < (L_{TJ})_{Md}^* = (L_{TJ})_e^* \), confirms that this is the case. \((L_{TJ})_n^* < (L_{TJ})_d^* < (L_{TJ})_e^* \) holds because in

---

34 The total perceived distance to the "average national" of modern sector workers could be greater than their distance to the "average co-ethnic", because of the greater ethnic distance to the "average national". However, the difference between the total distance under national identity and the one under ethnic identity of modern sector workers is always smaller than that of traditional sector workers, which implies that the incentive to identify with the nation is stronger for modern sector workers.

35 Robinson (2014) classifies workers into the formal and informal sectors based on their occupation: formal sector occupations are military/police, clerical worker, business person, professional worker, civil servant, teacher, etc., and informal sector occupations are subsistence farmer, informal manual labor, herder, housewife, etc.

36 By contrast, Eifert, Miguel, and Posner (2010), based on surveys in 10 African countries, find that being a farmer or fisherman, whom they classify as traditional sector workers, is negatively correlated with ethnic identity. However, there is no option for national identity in the surveys (other options are religious and class/occupational identities) and, unlike Robinson (2014), they classify those in the urban informal sector as modern sector workers.
equilibrium (d), the proportion of those identifying with the nation, who gain less from choosing the traditional sector, is higher (lower) than in equilibrium (e) (equilibrium (n)). \( (LT_J)_T^T = (LT_J)_n^T \) and \( (LT_J)_M^M = (LT_J)_e^e \) hold for the remaining heterogenous identity equilibria because individuals who identify with a particular identity are in both sectors (the national identity for equilibrium (Td) and the ethnic identity for equilibrium (Md)) and they are indifferent between the sectors, as in the corresponding homogenous identity equilibrium.

Finally, the result for the total output of private goods is similar to Proposition 1 and can be explained as before. Under plausible conditions, total output is generally higher as the proportion of individuals having the national identity is higher, although \( Y_{T_d}^* = Y_n^* \) and \( Y_{M_d}^* = Y_e^* \) are true. Further, because of the lower cost of conflict, aggregate material payoff is strictly higher when the proportion of those identifying with the nation is higher.

To summarize, results similar to Proposition 1 hold when heterogenous identity equilibria are also considered: prevalence of national identity is associated with not only a lower level of conflict but also higher shares of modern sector workers and production and, under plausible conditions, higher levels of the total output of private goods and of aggregate material payoff (the value of private and club good consumption net of the cost of conflict). The results are consistent with the often-made argument (Collier, 2009; Michalopoulos and Papaioannou, 2015) that the dominance of subnational (particularly ethnic) identities over national identity lies behind poor performance in various dimensions, including conflict and economic development, in ethnically heterogenous societies.\(^{37}\) Note that a small share of modern sector workers under widespread ethnic identity implies a large intersectoral gap in earnings. This suggests that strong ethnic identity might partly explain a substantial gap in average labor productivity between agriculture and non-agriculture in many developing countries found by Gollin, Lagakos, and Waugh (2014) and others.

### 3.3 Interactions among modernization, identity, conflict and output

The previous sections compared different equilibria for given parameters and exogenous variables. However, as indicated by Figures 1 and 2, the set of realized equilibria changes with values of exogenous variables. Taking this into account, this section analyzes the focus of the paper: the interactions among modernization, identity, conflict, and output.

A simple dynamics is introduced into the model by supposing that the TFP (total factor productivity) of sector \( M, A_M \), increases over time, i.e., \( A_{M,t+1} > A_{M,t} \) for any period \( t \). The TFP growth corresponds to the technological progress of the modern sector and the improvement in quality of institutions supporting the sector’s economic activities in the real economy. The productivity growth raises the modern sector income, induces a higher proportion of workers to choose the sector, i.e., lowers \( LT_J \), and raises the modern sector’s share in production. How does modernization driven by the productivity growth affect social identity, conflict, and aggregate

---

\(^{37}\)Michalopoulos and Papaioannou (2015) find a positive relationship between national identification and a measure of state capacity in protecting property rights and a positive relationship between ethnic identification and a measure of the inefficiency of the legal system, by using Afrobarometer Surveys covering 18–20 sub-Saharan nations.
output? The next proposition, based on the propositions in Appendix A, shows that the effect differs depending on the status difference \( \Delta S \equiv S_N - S_E \). Note that changes in exogenous variables such as a decrease in contested resources \( V \) have similar effects to an increase in \( \Delta S \), as shown later in Propositions 4 and 5.

**Proposition 3** Suppose that the TFP of sector \( M \), \( A_M \), increases over time.

(i) If the status difference \( \Delta S \) is very high (very low), the society always is in equilibrium (\( n \) (equilibrium \( (n) \))) and thus the level of conflict \( F \) is consistently low (high).

(ii) Otherwise, when \( \Delta S \) is relatively high (low), the society shifts from a heterogenous identity equilibrium to equilibrium \( (n) \) (equilibrium \( (e) \)), or stays in the latter equilibrium. When \( L_{TJ} \) is relatively large, given parameters and exogenous variables including \( \Delta S \), multiple equilibria exist and thus social identity, conflict, and output differ depending on which equilibrium is realized.

(iii) For a given \( A_M \), when \( \Delta S \) is high (low), the society is in an equilibrium with a high (low) proportion of people identifying with the nation and relatively low (high) \( F \), and the equilibrium is characterized by relatively low (high) \( L_{TJ} \) and unless \( \alpha \) or \( \beta \) is very high, relatively high (low) levels of \( Y \) and aggregate material payoff.

If the status difference \( \Delta S \) is at extreme levels, the society stays in the same equilibrium: when the national status compared with the ethnic status is very high (low), everyone always identifies with the nation (his or her ethnic group), and the level of conflict \( F \) is consistently low (high).

Otherwise, when \( \Delta S \) is relatively high (low) and the society starts with a heterogenous identity equilibrium, it shifts from the equilibrium, in which modern sector workers are more likely to have a national identity than traditional sector workers, to the one in which everyone identifies with the nation (their ethnic group) and the level of conflict is low (high). That is, the social identity initially associated with modern (traditional) sector workers eventually becomes the shared identity, when the status difference is relatively high (low).

Although the increase of the modern sector productivity, \( A_M \), always lowers \( L_{TJ} \) and raises the sector’s share in production, for a given \( A_M \) (i.e., given levels of modern sector technology and of the quality of institutions supporting the sector’s economic activities), the society with high (low) \( \Delta S \) achieves relatively large (small) modern sector shares in employment and production and, under plausible conditions, relatively high (low) levels of the aggregate output of private goods and aggregate material payoff. Hence, having sufficiently high national status relative to ethnic status is crucial for achieving good outcomes in development as well as in national identity and conflict.

However, history or "luck" is also important, as long as the status difference is not at extreme levels. For given parameters and exogenous variables including \( \Delta S \), multiple equilibria exist particularly when \( L_{TJ} \) is relatively large, thus, as shown in Propositions 1 and 2, social identity, conflict, conflict.

\[ \text{Note that modernization is not the same as urbanization: traditional sectors correspond to the urban informal sector as well as traditional agriculture and household production in the real economy. Many developing countries have experienced rapid urbanization without significant modernization.} \]

\[ \text{When } \Delta S \text{ is in the intermediate range, multiple equilibria exist for any } L_{TJ} \text{, and the society shifts from a heterogenous identity equilibrium to either equilibrium } (n) \text{, equilibrium } (e) \text{, or equilibrium } (Md). \]
and output differ depending on which equilibrium is realized. Suppose that an equilibrium realized initially is sustained in subsequent periods, as long as the equilibrium continues to exist. Then, if the initial equilibrium happens to be such that a relatively high proportion of individuals identify with the nation, the society tends to subsequently maintain a relatively strong national identity and relatively good outcomes in conflict, modern sector shares, and aggregate output.

3.3.1 Mechanism

The result would be understood more easily, by looking at the result when $\eta_1 = 0$ first, that is, when the weight on the ethnic attributes $\omega_e$ of the perceived distance does not depend on the level of conflict $F$ (see (11) in Section 2). In this case, $F$ is the same in all equilibria, and equilibrium is unique for given parameters and exogenous variables. Figure 3 illustrates how the realized equilibrium differs depending on $\Delta S$ and $L_{TJ}$. As $A_M$ increases over time, $L_{TJ}$ decreases, where its value is determined by the indifference condition for the sectoral choice of the corresponding equilibrium. In the figure, the society moves leftward with the productivity growth.

When the national status compared with the ethnic status is very high (very low), everyone always identifies with the nation (his or her ethnic group). By contrast, when the status difference is neither very high nor very low, the realized equilibrium changes with the productivity growth.

---

$^{40}$Multiple equilibria also exist in the model of Sambanis and Shyao (2013) when $\Delta S$ is not at extreme levels.

$^{41}$In particular, as noted in footnote 39 attached to the proposition, when $\Delta S$ is in the intermediate range, multiple equilibria exist for any $L_{TJ}$ and thus history or "luck" matters even in the long run. When $\Delta S$ is not in the intermediate range, the society shifts to a homogenous identity equilibrium and the effect of history or "luck" disappears eventually.

$^{42}$Positions of the dividing lines depend on parameters and several exogenous variables, but not on $A_T$ and $A_M$. 

---
When the status difference is relatively high (low), the society shifts from the equilibrium in which sector $M$ workers identify with the nation and sector $TJ$ workers identify with their ethnic group to the equilibrium in which all workers identify with the nation (their ethnic group). That is, the social identity initially associated with modern (traditional) sector workers eventually becomes the shared identity, when the status difference is high (low). The growth of $A_M$ raises the modern sector income and induces a higher proportion of workers to choose the sector. As a result, the national identity of modern sector workers becomes weaker in the sense that the difference between their utility and the utility when they switch to ethnic identity decreases, because the sectoral distance of these workers under ethnic identity falls more than the distance under national identity. Urban modern sector workers, who used to find little affinity with most of their ethnic group in rural areas or in the urban informal sector, feel closer to their group because there are more co-ethnics in occupations or with lifestyles similar to theirs. The ethnic identity of traditional sector workers also becomes weaker because a smaller proportion of their fellow group is in their sector. That is, the sectoral shift of labor associated with modernization shakes long-standing identities in both sectors.

When the national status relative to the ethnic status is high, the latter effect on traditional sector workers determines the equilibrium shift (because utilities under national identity are relatively high and thus modern sector workers are resilient to the "identity shock"), and all workers come to identify with the nation, while when the status difference is low, the former effect on modern sector workers determines the shift and all workers come to identify with their ethnic group.

When $\eta_1 > 0$, that is, when the weight on ethnic attributes $\omega_e$ of the perceived distance increases with $F$, $F$ is lower in an equilibrium with a higher proportion of individuals identifying with the nation (Propositions 1 and 2). Unlike the case $\eta_1 = 0$, multiple equilibria could exist for given parameters and exogenous variables, and two heterogenous identity equilibria $(Md)$ and $(Td)$ could also exist. Multiple equilibria could arise because of two-way positive causations between conflict and identity: when the level of conflict is high (low), people care about ethnicity more (less) in measuring the distance to social groups and thus they are more (less) likely to identify with their ethnic group, whereas when the proportion of those with an ethnic identity, who do not care about the distance to other ethnic groups, is high (low), the level of conflict is high (low). Equilibria $(Md)$ and $(Td)$ could exist because the level of conflict depends on identity when $\eta_1 > 0$: workers in one of the sectors can be indifferent between the two identities only if their identity choice affects $f_i$ and thus their utilities (the indifference conditions (29) and (32) do not hold when $\eta_1 = 0$).

Figure 4, which is essentially the same as Figure 2, illustrates how realized equilibria differ depending on $\Delta S$ and $L_{TJ}$ when $\eta_1 > 0$. Equilibrium $(n)$ (equilibrium $(e)$) exists in the region above the upward-sloping solid curve (on or below the downward-sloping solid curve). Equilibrium $(d)$ exists in the region with triple-dashed double-dotted lines, equilibrium $(Md)$ exists in the region with negatively-sloped lines, and equilibrium $(Td)$ exists in the region with positively-sloped lines.\footnote{\textit{The sectoral distance of modern sector workers falls under either identity, because a higher proportion of workers is in the sector. However, the fall under ethnic identity is greater because changes in the average sectoral attributes are greater (see (15) and (21)).}}

\footnote{\textit{Remember that equilibrium $(d)$ is the equilibrium in which sector $M$ workers identify with the nation and sector}}
Suppose, for example, that the society starts with equilibrium (d). As long as it stays in this equilibrium, with the growth of $A_M$, $L_{TJ}$, $L_T$, and thus the proportion of individuals with an ethnic identity decrease, which leads to a fall in the level of conflict as well as to increases in the modern sector shares of employment and production, output, and material payoff. Eventually, this equilibrium ceases to exist, and the society shifts to a different equilibrium. If $\Delta S$ is relatively high (low), it shifts to equilibrium (n) (equilibrium (e)) and the level of conflict falls (rises).45 (The rise of $F$ when $\Delta S$ is low may be interpreted as a rise in non-violent conflict such as rent-seeking activities, if the shift occurs at a relatively low $L_{TJ}$, i.e., at a later stage of economic development.) The sectoral shift and output growth continue, but for a given level of the modern sector productivity, $L_{TJ}$ is lower, and the modern sector shares, output, and material payoff are higher in the equilibrium of universal national identity.

The figure shows that there are several regions in which multiple equilibria exist. For example, suppose that the society starts with the region in which equilibria (e), (d), and (Md) exist. Depending on which equilibrium happens to be realized initially, social identity, the level of conflict, sectoral composition, and aggregate output differ in initial and subsequent periods (even in the long run when $\Delta S$ is in the intermediate range): the outcome is worst when the society starts with equilibrium (e) and is best when it starts with equilibrium (d).

$TJ$ workers identify with their ethnic group, and equilibrium (Md) (equilibrium (Td)) is the one in which sector $M$ workers are divided over identities and sector $TJ$ workers share the ethnic identity (sector $TJ$ workers are divided over identities and sector $M$ workers share the national identity).

45 The equilibrium shift changes $F$ discontinuously, while it changes $L_{TJ}$ continuously. When $\Delta S$ is in the intermediate range, as shown in the figure and mentioned in footnote 39, the society shifts to either equilibrium (n), equilibrium (e), or equilibrium (Md).
3.3.2 Discussion

As mentioned in the Introduction, competing theses exist in political science on the effects of modernization on social identity. The classic thesis, which is based on the past experience of Europe, argues that modernization leads to widespread national identity at the expense of ethnic and other subnational identities (Deutsch, 1953; Gellner, 1964, 1983; Weber, 1979), while another influential thesis ("second-generation" thesis) mainly focusing on Africa argues that modernization rather breeds ethnic identification (Melson and Wolpe, 1970; Bates, 1983).

Proposition 3 shows that when the national status compared with the ethnic status is relatively high (low) and the society starts with a heterogenous identity equilibrium, the society shifts to the equilibrium with universal national (ethnic) identity that is characterized by a low (high) level of conflict, high (low) modern sector shares of employment and production, and high (low) aggregate output and material payoff. Thus, the result is consistent with the classic view when the status difference is relatively high, while it is consistent with the "second-generation" view when it is relatively low, as far as the relatively long-term effect of modernization (the effect involving the equilibrium shift) is concerned.

Robinson (2014), by contrast, using cross-sectional individual-level survey data of sixteen African nations, finds that GDP per capita is significantly and positively related to individuals identifying more with the nation than with their ethnic group, after controlling for various individual-level (such as formal sector employment), group-level and country-level variables. She interprets the evidence as suggesting that modernization (higher GDP per capita) leads to national identity. Considering that it is based on cross-sectional data of mostly poor African nations, the evidence may be regarded as capturing the relatively short-term effect in an economy with a low degree of modernization. Indeed, the effect of an increase in $A_M$ under a given equilibrium is consistent with her interpretation, when $\Delta S$ is not at extreme levels and $L_{TJ}$ is sufficiently high that the society is in a heterogenous identity equilibrium (except equilibrium $(T_d)$, in which the effect is negative). The evidence can also be interpreted differently, however, and certain results of the model are consistent with the alternative interpretations.

Proposition 3 also shows that multiple equilibria exist and thus the outcome depends on history or "luck" when the status difference is not at extreme levels. As Sambanis and Shayo (2013) stress,

\[ F \text{ increases with a decrease in } L_{TJ} \text{ from (32)}. \]
\[ \text{Then, the number of those identifying with their ethnic group, } (1-P_{TJ,n})n_eL_{TJ}, \text{ increases with a decrease in } L_{TJ} \text{ from (33)}. \]

The evidence is partly consistent with the story that for a given modern sector productivity $A_M$, national identity and modernization are positively related through positive effects of the status difference on these variables (Figure 4 and Proposition 3 (iii)). That is, when $\Delta S$ is higher, for a given $A_M$, the society is in an equilibrium with a higher proportion of individuals identifying with the nation and a lower $L_{TJ}$ (thus a higher degree of modernization). Further, the evidence can also partly be explained by multiple equilibria, because for given parameters and exogenous variables, national identity and modernization are positively related among different equilibria (Propositions 1 and 2). To distinguish the different stories empirically, it would be important to estimate regression models with enough control variables (including measures capturing the status difference and the modern sector productivity) using longitudinal data, although such data is not available presently. Analysis using longitudinal data is also called for to examine empirically the relatively long-term effect of modernization, on which this paper and the above-cited studies in political science focus.
this is consistent with the evidence that countries similar in ethnic diversity, geography, economic conditions, and political institutions have diverse histories regarding levels of ethnic conflict.

What are the policy implications of the finding that having sufficiently high national status relative to ethnic status is crucial for good outcomes? Miguel (2004), Collier (2009), and Blouin and Mukand (2019), based on a case study or statistical analysis, argue that nation-building policies, such as school education or government propaganda emphasizing common history, culture, and values and the promotion of a national language, are effective in strengthening national identity. Classic modernization theories of nationalism also stress the importance of the unification of language and the spread of common culture and values through school education and universal military service for widespread national identity, by drawing on the experience of Europe.

Some of these nation-building policies would raise pride of the people in the nation and thus may be interpreted as policies raising $S_N$ and $\Delta S$. The model shows how they can reinforce national identity and produce good outcomes in conflict and development when $\Delta S$ is not high. The result suggests that these policies are critical for achieving good outcomes in countries such as many African nations, where the perceived national status is low in comparison with ethnic status because of the lack of shared culture, history, and values. Without nation-building policies, policies promoting modernization (policies raising $A_M$) have negative effects on national identity and ethnic conflict and are not very effective for development in these countries. Ethnic-based policies, such as education emphasizing unique history, culture, and values of each ethnic group and the promotion of ethnic languages, would also have negative effects on national identity, conflict, and development, because they can be considered as policies raising $S_E$ and lowering $\Delta S$. By contrast, the result suggests that in countries with high perceived national status, policies promoting modernization, such as policies stimulating the technological progress of the modern sector, the reform of institutions supporting the sector’s economic activities, and the construction of transportation infrastructure connecting rural areas to urban areas, would be enough to achieve good outcomes.

Alternatively, some of the nation-building policies, particularly the promotion of a national language and education or propaganda emphasizing common values, may be interpreted as policies making ethnic differences less salient, i.e., reducing $\omega$, in the perceived distance. Indeed, the following proposition shows that a decrease in $\eta_0$ that lowers $\omega$ has similar effects to an increase in $\Delta S$.

**Proposition 4** Suppose that $A_M$ increases over time. Then, results similar to Proposition 3 hold for $\eta_0$ when a "high (low) $\Delta S$" is replaced with a "low (high) $\eta_0$". Specifically, given other things being equal, when $\eta_0$ is low (high), the society shifts from a heterogenous identity equilibrium to equilibrium (n) (equilibrium (e)), or stays in the latter equilibrium.

---

48 By contrast, Alesina, Giuliano, and Reich (2018) model nation-building policy as the policy that homogenizes the cost of the preference distance from the government of individuals with heterogenous preferences.

49 The effect of transportation infrastructure can be examined in the slightly modified model in which earnings in the traditional sector are lower than earnings in the modern sector partly due to the presence of the cost of moving from the former sector to the latter. Improved infrastructure lowers the moving cost and stimulates modernization.
Graphically, this result holds because all the dividing lines of Figure 4 shift downward when $\eta_0$ decreases.

The importance of nation-building policies on social identity seems to be supported empirically. Miguel (2004) bases the above mentioned argument on findings from the comparison of Tanzania and Kenya, which largely share geography, history, and a colonial institutional legacy, but conducted sharply different ethnic policies after independence in areas, such as national language and public school education, and exhibit large differences in the strength of their national identity. Clots-Figueras and Masella (2013) examine the effects of the introduction of a bilingual (Catalan and Spanish) education system in Catalonia and find that the number of years exposed to the bilingual system is positively related to the strength of the Catalan identity and the propensity to vote for a party with a Catalanist platform. Blouin and Mukand (2019), based on field and lab experiments in post-genocide Rwanda, find that ethnicity is less salient in regions exposed more to government radio propaganda intended to change inter-ethnic attitudes.

### 3.4 Effects of contested resources

Finally, the effects of the amount of contested resources are examined. Specifically, how does it affect the level of conflict, and how does it influence the effects of modernization on identity, conflict, and development?

**Proposition 5**

(i) The level of conflict $F$ increases with resources $V$ in all equilibria.

(ii) Suppose that the TFP of sector $M$, $A_M$, increases over time. Then, results similar to Proposition 3 hold for $V$ when a "low (high) $\Delta S$" is replaced with a "large (small) $V$". Specifically, when $V$ is large (small), the society shifts from a heterogenous identity equilibrium to equilibrium $(e)$ (equilibrium $(n)$), or stays in the latter equilibrium.

The first result, which is consistent with empirical studies on armed internal conflict such as Collier and Hoeffler (2004), is standard and intuitive: as the amount of contested resources increases, people contribute more to conflict, and thus the conflict level increases.

The second result states that similar results to Proposition 3 hold when the "low (high) $\Delta S$" of the proposition is replaced with a "large (small) $V$". Specifically, *given the status difference*, when the amount of contested resources is large (small), with an increase in modern sector productivity, the society shifts from a heterogenous identity equilibrium to equilibrium $(e)$ (equilibrium $(n)$), or stays in the latter equilibrium. Hence, the abundance of contested resources is an impediment for the society in achieving universal national identity, a low level of conflict, high modern sector shares, and high output. Note that contested resources represent not only material resources (e.g., natural resources) but also a part of the governmental budget that can be used to benefit particular ethnic groups, whose allocations among the groups are determined not by rule but by the consequences of violent or non-violent conflict (e.g., rent-seeking activities). Hence, the result suggests that not only the abundance of resources per se but also the lack of strong political and economic institutions (e.g., weak rule of law), which leads to abundant contested resources, are
hindrances to a desirable outcome. The result is consistent with the classic thesis on the effects of modernization on identity if resources are not abundant or the institutions are good in quality, otherwise, it is consistent with the competing thesis, as far as the relatively long-term effect of modernization is concerned.

Graphically, the result holds because all the dividing lines of Figure 4 shift upward when $V$ increases. When the amount of contested resources increases, the level of conflict rises in all equilibria, and thus people care about ethnicity more, i.e., $\omega_e$ increases, in measuring perceived distances to social groups. Given $\Delta S$, this makes identifying with their ethnic group relatively more attractive for people than identifying with the nation.

Consistent with the result, Mehlum, Moene, and Torvik (2006) find negative effects of natural resources on economic development when institutions are weak. Empirical works also suggest that political and economic institutions have important effects on civil conflict (Renyal-Querol, 2002), rent-seeking activities (Easterly, 2001), and development (Rodrik, Subramanian, and Trebbi, 2004). The above result reveals a novel mechanism interacting with social identity by which resources and institutions affect ethnic conflict and development.

In the real society, modernization often increases the amount of contested resources, such as public-sector jobs, approvals and licenses for regulated business activities, and budgets for local schools and infrastructures, when political and economic institutions are weak. The above result implies that, as "second-generation" modernization theories of nationalism (Melson and Wolpe, 1970; Bates, 1983) argue, by raising the amount of contested resources, modernization has negative effects on national identity, ethnic conflict, and development in a society with low institutional quality. The result in Section 3.3, however, shows that even when contested resources do not increase with modernization, modernization could lead to negative outcomes in a society with low institutional quality, low national status relative to ethnic status, or people who care greatly about ethnic differences.

3.5 Discussions on several assumptions

The model has imposed several assumptions that make analysis tractable. This section briefly discusses how results would be affected and what kind of new questions could possibly be examined when these assumptions are replaced by more realistic ones.

3.5.1 Assumption on benefit of ethnic-specific club goods

In the model, individuals participate in ethnic conflict even when they have a national identity or are in the modern sector. This might be justified considering the fact that resources are contested violently or non-violently between ethnic groups, not between socioeconomic classes or groups with different ideologies, in many developing countries, particularly in sub-Saharan African countries (Posner, 2005). That is, even when individuals identify with the nation or are in the relatively ethnically-integrated modern sector, they have no choice but to engage in ethnic conflict to receive a share of contested resources in many countries.
However, for those with a national identity, the benefit of some ethnic-specific club goods, such as public spending on ethnic culture and language, might be small, while the benefit of some public goods, such as spending on common culture and language, might be large. How are results affected when this factor is taken into account?

Suppose that a fixed amount of resources exists. A part of the resources are allocated to the government based on rules and are used, for simplicity, entirely for the provision of public goods, while the rest of the resources $V$ are allocated between ethnic groups as a result of ethnic conflict and are used for the provision of ethnic-specific club goods. The proportion of the resources allocated for each use is assumed to be fixed for simplicity. This might be justifiable because the allocation largely depends on the quality of political institutions presumed to be constant in the model: good institutions inhibit rent-seeking activities and violent conflict and thus lower the resources used for ethnic-specific goods. Assume that the preference for group-specific club goods is stronger for individuals with an ethnic identity than for individuals with a national identity, i.e., $\delta_e > \delta_n$, while the benefit of public goods is normalized to be 0 for those with an ethnic identity and is positive for those with a national identity.

Assuming $\delta_e > \delta_n$ increases the difference in contribution to conflict $f_i$ between those with an ethnic identity and those with a national identity and in the level of conflict $F$ between an equilibrium with a large number of people having an ethnic identity and the one with a large number of people having a national identity. However, qualitative results are not affected.

Alternatively, the benefit of ethnic-specific club goods, such as education, health services, and roads for particular groups or for areas they are clustered, could be small for modern sector workers, because they work and live in relatively ethnically integrated environments, whereas the benefit of public goods, such as legal service and scientific knowledge, could be large for these workers. This factor can be taken into account by assuming that $T > M$ and the benefit of public goods is normalized to be 0 for traditional sector workers and is positive for modern sector workers.

This modification does not affect the main result that the society ends up with an equilibrium of universal ethnic (national) identity when the status difference $\Delta S$ is small (large) or when $\omega_e$ in the perceived distance or the amount of contested resources $V$ is large (small). However, several results are qualitatively affected. In particular, because of $\delta_T > \delta_M$, modern sector workers contribute less to conflict than traditional sector workers even when they share the same identity. Hence, an increase in $A_M$ lowers the level of conflict $F$ when the society is in a homogenous identity equilibrium. Under the original setting, when $\Delta S$ is relatively but not extremely small (when $\omega_e$ or $V$ is relatively but not extremely large), as the society shifts from a heterogeneous identity equilibrium to equilibrium (e) with an increase in $A_M$, $F$ increases. By contrast, when $\delta_M$ is sufficiently small relative to $\delta_T$, modernization always lowers the level of conflict.

---

$^{50}$ $f_i = 0$ does not occur even for those with a national identity unless $\delta_n = 0$. This is because the cost of contributing to conflict is given by $c(f_i) = \frac{1}{2}(f_i)^\theta$, $\theta \geq 2$, where the assumption $\theta \geq 2$ follows Esteban and Ray (2011) and is needed to prove some results. If $\theta = 1$ is assumed as in Sambanis and Shayo (2013), $f_i = 0$ can occur.

$^{51}$ Regarding heterogeneous identity equilibria, as under the original setting, an increase in $A_M$ lowers (raises) $F$ in equilibria (d) and (Md) (in equilibrium (Td)).
3.5.2 Assumption of symmetric ethnic groups

In the model, ethnic groups are assumed to be symmetric in every respect for analytical tractability and for the focus on a society without a dominant ethnic group. In the real world, of course, there are many societies in which dominant groups exist. Hence, it would be important to discuss what kind of new questions could possibly be examined using extended models with asymmetric ethnic groups, although analysis of such models is very complicated.

One interesting question arising in a society with asymmetric ethnic groups is whether a large ethnic group is more likely to identify with the nation than a small group or not. This question could be examined using a simplest model with asymmetric groups in which the population of group 1 is much larger than that of each of groups 2 to $n_e$ and the smaller groups are symmetric in every respect. It would be realistic to suppose that, because of the much greater population, the status of group 1 is higher than that of the other groups. This implies that given $S_N$, $\Delta S$ is smaller for group 1, which makes the group less likely to identify with the nation than the other groups. At the same time, the perceived distance (both the ethnic distance and the sectoral distance) under national identity is smaller for group 1 because of their population size, which makes the group more likely to identify with the nation. Thus, group 1 individuals are more likely to have an ethnic identity than the other groups when the status of group 1 is sufficiently higher than the other groups, otherwise, they are more likely to have a national identity. In particular, when $S_N$ is relatively but not extremely small, as modernization proceeds, the society would shift to an equilibrium with a universal ethnic identity for group 1 and a universal national identity for the smaller groups in the former case, whereas in the latter case, the society would shift to an equilibrium with a universal national identity for group 1 and universal ethnic identities for the other groups.

A model with asymmetric group could also be used to examine the relationship between measures of divisions or differences of ethnic groups and conflict. Esteban, Mayoral, and Ray (2012) show empirically that measures of the intensity of civil conflict are significantly related to three measures of ethnic divisions, positively to the polarization index and the fractionalization index, and negatively to the Greenberg-Gini index, and the effect of polarization (fractionalization) on conflict increases (decreases) when public or club goods rather than private goods become more important as rewards to conflict.52 Their analysis is based on the theoretical study by Esteban and Ray (2011), which show that, in a contest model of conflict (without socio-psychological factors), the equilibrium level of conflict is an increasing function of a linear combination of the three measures of ethnic divisions. In particular, they show that, when ethnic groups compete solely for resources for group-specific club goods as in the present model, the level of conflict depends on the polarization index and the Greenberg-Gini index.

52 The polarization index is $P = \sum_{i=1}^{n_e} \sum_{j=1}^{n_e} \left( \frac{L_i}{L_j} \right)^2 \frac{L_j}{2} \delta_{ij}$, the fractionalization index is $F = \sum_{i=1}^{n_e} \frac{L_i}{2} \left( 1 - \frac{L_i}{2} \right)$, and the Greenberg-Gini index is $G = \sum_{i=1}^{n_e} \sum_{j=1}^{n_e} \left( \frac{L_i}{L_j} \right) \frac{L_j}{2} \delta_{ij}$, where $\delta_{ij}$ is the difference in the benefits for group $i$ individuals of their own club goods and of club goods of group $j$. In the present model, $\delta_{ij} = V$ for $i \neq j$ (because the benefits of club goods of other groups are 0) and $\delta_{ij} = 0$ for $i = j$. 

30
An examination of the present model of symmetric ethnic groups, in which values of the three measures are determined by the number of ethnic groups \( n_e \), reveals that in homogenous identity equilibria, the result of Esteban and Ray (2011) applies, although in order for the polarization index to have influence on conflict, as shown in Esteban and Ray (2011), individual utility must also depend on the utilities of other members of the group he or she identifies with. In heterogeneous identity equilibria, by contrast, \( F \) depends on \( n_e \) not only through the measures of ethnic divisions but also through variables related to population sizes of each ethnic group and of each sector. In an extended model with asymmetric groups, the result of Esteban and Ray (2011) would hold when all ethnic groups are in the same homogenous identity equilibrium, while in the remaining cases, \( F \) would depend on the demographic variables as well as on the measures of ethnic divisions.

4 Conclusion

Empirical evidence suggests that ethnic divisions in a society lead to negative outcomes in various dimensions, including civil conflict and economic development. It is often argued that the lack of shared social identity lies behind the negative outcomes in ethnically heterogenous societies. If shared national identity is important, how can it be realized? In political science, conflicting theses exist regarding the effects of modernization on national identity, the classic thesis claiming the positive effect and the competing one claiming the negative effect. Which thesis is more relevant under what conditions? How does modernization affect identity, conflict, and output? Do nation-building policies, such as school education or government propaganda emphasizing common history, culture, and values and the promotion of a national language, influence the consequences of modernization for identity and other outcomes?

To examine these questions theoretically, this paper has developed a model of social identity, ethnic conflict, and development. In the model, individuals choose a sector to work for (between the modern sector and a traditional sector), a social identity (between ethnic identity and national identity), and a contribution to conflict. The degree of modernization (and output), identity, and conflict interact each other in the model.

It has been found that a society with people having stronger pride in the nation or caring less about ethnic differences, less contested resources, or better institutions is in an equilibrium with a higher proportion of people identifying with the nation, a lower level of conflict, a higher modern sector share in employment and production (a higher degree of modernization), and higher output. A simple dynamic analysis has shown that, as modernization proceeds, a society shifts to an equilibrium with a universal national identity and good outcomes in conflict and development, if the national pride is high, ethnic differences are not salient in people’s minds, the contested resources are not abundant, or institutions are good in quality; otherwise, it shifts to an equilibrium with a universal ethnic identity and worse outcomes in other dimensions. Hence, the result is consistent

\[ 53 \] In the present model, values of the fractionalization index and the Greenberg-Gini index equal \( \left(1 - \frac{1}{n_e}\right)V \) and the value of the polarization index equals \( \frac{1}{n_e} \left(1 - \frac{1}{n_e}\right)V \), which is decreasing in \( n_e \).
with the classic (competing) thesis on the effects of modernization on identity under the former (latter) situation. The result suggests that under the latter situation, policies improving institutional quality, raising the national pride, or making ethnic differences less salient are crucial for achieving good outcomes. The model shows how nation-building policies, which may be interpreted as policies raising the national pride or making ethnic differences less pronounced, can reinforce national identity and produce better outcomes. The model also has revealed a novel mechanism interacting with identity by which resources and institutions affect conflict and development.

References


[27] Ezcurra, Roberto Ezcurra and Andrés Rodríguez-Pose (2017), "Does ethnic segregation matter for spatial inequality?," *Journal of Economic Geography* 17, 1149–1178.


[34] Lewis, A. W. (1954), "Economic development with unlimited supplies of labour," Manchester School of Economic and Social Studies 22 (2), 139–191.


Appendix A  Existence conditions of equilibria

This Appendix presents precise conditions (combinations of parameters and exogenous variables) under which each equilibrium exists. The propositions in this Appendix are the basis for Propositions 3–5 and all figures in Section 3.

A.1 Homogenous identity equilibria

The next proposition presents the existence conditions for the two homogenous identity equilibria. In the proposition, \( \beta \Delta d^2[F,c_s] \equiv \beta \left[ \eta_0 + \eta_1 F \right] \frac{n_e - 1}{n_e} + c_s \omega_s \) \( (c_s \text{ is a coefficient on } \omega_s) \), \( \Delta c(F) \equiv \frac{1}{2} \left( \frac{n_e - 1}{n_e} \right)^{\frac{1}{T}} \left[ \left( V_T^{\frac{1}{T}} - (\delta V_T - \beta \eta)^{\frac{1}{T}} \right) \right] \), and \( \Delta S \equiv S_N - S_E \).

Proposition A1

(i) Equilibrium \((e)\) exists for any \( L_{TJ} \) when \( \gamma \Delta S \leq \beta \Delta d^2 \left[ F_e^*, -\frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*) \), and for \( L_{TJ} \in [0, (L_{TJ})^{\dagger\dagger}] \) when \( \gamma \Delta S \in \left( \beta \Delta d^2 \left[ F_e^*, -\frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*) \right), \beta \Delta d^2 \left[ F_e^*, 0 \right] - \Delta c(F_e^*) \), where \( L_{TJ} = (L_{TJ})_e^* \) is the solution for (20) and \( (L_{TJ})^{\dagger\dagger} \) is the one for \( \beta \Delta d^2 \left[ F_e^*, -n_e(n_e - 1) \left( \left( \frac{K_T}{T} \right)^{\dagger\dagger} \right) \right] - \Delta c(F_e^*) = \gamma \Delta S \).

\(^{54}(L_{TJ})^{\dagger\dagger} (L_{TJ})^{\dagger\dagger}\) is the downward-sloping (upward-sloping) curve of Figure 1.
(ii) Equilibrium (n) exists for any \( L_{TJ} \) when \( \gamma \Delta S > \beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) \), and for \( L_{TJ} \in [0, (L_{TJ})^a] \) when \( \gamma \Delta S \in \left( \beta \Delta d^2[F_n^*, 0] - \Delta c(F_n^*), \beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) \right) \), where \( L_{TJ} = (L_{TJ})^a \) is the solution for (25) and \( (L_{TJ})^a \) is the one for \( \beta \Delta d^2 \left[ F_n^*, (n_e - 1) \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_n^*) = \gamma \Delta S \).

A.2 Heterogenous identity equilibria

The next proposition presents the existence conditions for the heterogenous identity equilibria.

Proposition A2 (i) Equilibrium (d) exists iff \( \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1)n_e \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_d(L_{TJ})) < \gamma \Delta S < \beta \Delta d^2 \left[ F_e^*, -(n_e - 1) \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_e^*) \), where \( F_d(L_{TJ}) \) is the solution for (26) and increases with \( L_{TJ} \) and \( L_{TJ} = (L_{TJ})^a \) is the solution for (27).\(^{55}\)

(ii) Equilibrium (Md) exists iff \( \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1)n_e \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_d(L_{TJ})) < \gamma \Delta S < \beta \Delta d^2 \left[ F_e^*, -(n_e - 1) \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_e^*) \), where \( L_{TJ} = (L_{TJ})^e \) is the solution for (20).\(^{56}\)

(iii) Equilibrium (Td) exists for \( L_{TJ} \in ([L_{TJ}]^b, (L_{TJ})^b) \) when \( \gamma \Delta S \in \left( \beta \Delta d^2[F_n^*, 0] - \Delta c(F_n^*), \beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) \right) \), and for \( L_{TJ} \in ([L_{TJ}]^b, \frac{L_{TJ}}{n_e}] \) when \( \gamma \Delta S \in \left( \beta \Delta d^2[F_n^*, \frac{n_e-1}{n_e}] - \Delta c(F_n^*), \beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) \right) \), where \( L_{TJ} = (L_{TJ})^b \) is the solution for (25), \( (L_{TJ})^b \) is the one for \( \beta \Delta d^2 \left[ F_d(L_{TJ})^b, -(n_e - 1) \left( \frac{L_{TJ}}{L} \right)^b \right] - \Delta c(F_d(L_{TJ})^b) = \gamma \Delta S \) and \( (L_{TJ})^b \) is the one for \( \beta \Delta d^2 \left[ F_e^*, (n_e - 1) \left( \frac{L_{TJ}}{L} \right)^b \right] - \Delta c(F_e^*) = \gamma \Delta S \).

Figure 2 in Section 3.2 illustrates combinations of \( L_{TJ} \) and \( \Delta S \) such that each equilibrium exists when \( \omega_i \) is relatively high and \( \eta_1 \) is relatively low.\(^{57}\) By contrast, Figure 5 illustrates combinations

\(^{55}\)To be more detailed, the equilibrium exists for \( L_{TJ} \in ([L_{TJ}]^a, \frac{L_{TJ}}{n_e}] \) when
\[ 
\gamma \Delta S \in \left( \max_{L_{TJ}} \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1)n_e \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_d(L_{TJ})) \right), \beta \Delta d^2 \left[ F_e^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_e^*) \text{ and for } L_{TJ} \geq \max \{ (L_{TJ})^a, 0 \} \text{ satisfying } \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1)n_e \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_d(L_{TJ})) < \gamma \Delta S \text{ when } \gamma \Delta S \in \left( \min_{L_{TJ}} \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1)n_e \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_d(L_{TJ})) \right). \]

\(^{56}\)To be more detailed, the equilibrium exists for \( L_{TJ} \in [0, (L_{TJ})^a] \) when
\[ 
\gamma \Delta S \in \left( \max_{L_{TJ}} \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1)n_e \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_d(L_{TJ})) \right), \beta \Delta d^2 \left[ F_e^*, 0 \right] - \Delta c(F_e^*) \text{ and for } L_{TJ} < \min \{ (L_{TJ})^a, \frac{L_{TJ}}{n_e}] \text{ satisfying } \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1)n_e \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_d(L_{TJ})) < \gamma \Delta S \text{ when } \gamma \Delta S \in \left( \min_{L_{TJ}} \beta \Delta d^2 \left[ F_d^*, 0 \right] - \Delta c(F_d^*) \right), \beta \Delta d^2 \left[ F_e^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_e^*) \text{ and for } L_{TJ} < \min \{ (L_{TJ})^a, \frac{L_{TJ}}{n_e}] \text{ satisfying } \beta \Delta d^2 \left[ F_e^*, -(n_e - 1)n_e \left( \frac{L_{TJ}}{L} \right)^a \right] - \Delta c(F_e(L_{TJ})) < \gamma \Delta S. \]

\(^{57}\)To be more accurate, Figure 2 illustrates the case when \( \beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) > \beta \Delta d^2 \left[ F_n^*, 0 \right] - \Delta c(F_n^*) \Rightarrow \beta \Delta d^2 \left[ F_n^*, 0 \right] - \Delta c(F_n^*) > \beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) \Rightarrow \beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) > \beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) + [\Delta c(F_n^*) - \Delta c(F_n^*)] \) holds, where \( \beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) \) is the value of \( \gamma S \) at the intersection of the upward-sloping (downward-sloping) solid curve with \( L_{TJ} = \frac{L}{n_e} \) \( (L_{TJ} = 0) \), and \( \beta \Delta d^2 \left[ F_n^*, 0 \right] - \Delta c(F_n^*) \) \( (\beta \Delta d^2 \left[ F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) \)
of $LTJ$ and $\Delta S$ when $\omega_s$ is relatively low and $\eta_1$ is relatively high. Unlike Figure 2, the value of $\Delta S$ at the intersection of the downward-sloping solid curve with $LTJ = 0$ is greater than the one at the intersection of the upward-sloping solid curve with $LTJ = \frac{n_e}{L}$, and the value of $\Delta S$ at the intersection of the bottom dotted curve with $LTJ = 0$ is smaller than the one at the intersection of the curve with $LTJ = \frac{n_e}{L}$. However, basic features of the figure are similar to Figure 2.

**Appendix B  Proofs (Perhaps not for publication)**

**Proof of the uniqueness of $(LTJ)^*_d$.** The derivative of the LHS of (27) with respect to $LTJ$ equals

\[-(1-\alpha)A_T(LTJ)^{\alpha-2} + 2\beta\omega_s \frac{n_e}{L} \left[ 1 - (n_e-1)\frac{LTJ}{L} \right] \]

\[+ \frac{n_e-1}{n_e} \left\{ \beta \eta_1 + \frac{1}{\frac{L}{d}((LTJ))} \left( \frac{n_e-1}{n_e} \right)^{\frac{1}{\frac{L}{d}}} \left[ \left( \frac{L}{d} \frac{V}{F_d(LTJ)} \right)^{\frac{1}{\frac{L}{d}}} - \left( \frac{L}{d} \frac{V}{F_d(LTJ)} - \beta \eta_1 \right)^{\frac{1}{\frac{L}{d}}} \right] \right\} F_d'(LTJ), \quad (34)\]

where, from (26),

\[F_d'(LTJ) = \frac{\left( \frac{n_e-1}{n_e} \right)^{\frac{1}{\frac{L}{d}}} \left[ \left( \frac{L}{d} \frac{V}{F_d(LTJ)} \right)^{\frac{1}{\frac{L}{d}}} - \left( \frac{L}{d} \frac{V}{F_d(LTJ)} - \beta \eta_1 \right)^{\frac{1}{\frac{L}{d}}} \right] n_e}{1 + \frac{1}{\frac{L}{d}} \frac{V}{F_d(LTJ)} \left( \frac{n_e-1}{n_e} \right)^{\frac{1}{\frac{L}{d}}} \left[ \left( \frac{L}{d} \frac{V}{F_d(LTJ)} \right)^{\frac{1}{\frac{L}{d}}} - \left( \frac{L}{d} \frac{V}{F_d(LTJ)} - \beta \eta_1 \right)^{\frac{1}{\frac{L}{d}}} \right]^{-1} n_e LTJ + \left( \frac{L}{d} \frac{V}{F_d(LTJ)} - \beta \eta_1 \right)^{\frac{1}{\frac{L}{d}}}^{-1} \left( L - n_e LTJ \right)} > 0. \quad (35)\]

The second derivative of the LHS of (27) with respect to $LTJ$ equals

\[\gamma S_N \text{ at the intersection of the bottom dotted curve with } LTJ = 0 \left( LTJ = \frac{L}{n_e} \right). \] The LHS of the last inequality increases with $\omega_s$, while the RHS can be shown to increase with $\eta_1$. 

37
\[(2-a)(1-a)A_T(L_{TJ})^{\alpha-3} - 2\beta\omega_s \frac{n_e(n_e-1)}{L^2}
\]
\[
d + \left( \frac{n_e-1}{n_e} \left[ \beta \eta_1 + \frac{1}{\theta - 1} \left( \frac{V}{(F_d(L_{TJ}))^2} \frac{(n_e-1)}{n_e} \right) \right] \right) \frac{1}{2} \left( \frac{\delta}{F_d(L_{TJ})} \right) \frac{1}{L^2} \left( \frac{\delta}{F_d(L_{TJ})} - \beta \eta_1 \right) \left( \frac{1}{L^2} \right) \frac{dL_{TJ}}{dL_{TJ}}. \tag{36}\]

Since, as shown in the proof of Proposition 2 (iii), the LHS of (27) is always lower than that of (20), \((L_{TJ})_d^* < (L_{TJ})_e^*\) holds. Thus, the derivative of the LHS of (20) for \(L_{TJ} \leq (L_{TJ})_d^*\) is negative, that is, \(-(1-a)A_T(L_{TJ})^{\alpha-2} + 2\beta\omega_s \frac{n_e}{L} < 0\). From this inequality, the first part of (36) is positive for \(L_{TJ} \leq (L_{TJ})_d^*\), since it is greater than \((1-a)A_T(L_{TJ})^{\alpha-3} (2-a) - L_{TJ} \frac{n_e}{L} > 0\).

The second part of (36) is positive at \(L_{TJ} \leq (L_{TJ})_d^*\) when \(\theta = 2\), since the second part of (34) equals
\[
\frac{n_e-1}{n_e} \beta \eta_1 \left[ 1 + \delta \left( \frac{V}{(F_d(L_{TJ}))^2} \right) \right] \frac{n_e-1}{n_e} \beta \eta_1 n_e \frac{1}{1+\delta \left( \frac{V}{(F_d(L_{TJ}))^2} \right) n_e-1} L, \tag{37}\]
which clearly increases with \(L_{TJ}\). Hence, the second derivative of the LHS of (27) is positive for \(L_{TJ} \leq (L_{TJ})_d^*\) and thus \((L_{TJ})_d^*\) is unique when \(\theta = 2\) (remember that the LHS of (27) is always lower than that of (20) whose solution is unique \((L_{TJ})_e^*\)).

**Proof of Proposition 1.** (i) Evident from (18) and (24). (ii) Shown in footnotes 23 and 24.

(iii) Denote the total output of private goods in equilibrium (n) (equilibrium (e)) by \(Y_n^* (Y_e^*)\). From (1) and (2),
\[
Y_n^* > Y_e^* \Leftrightarrow A_T((L_{TJ})_n)^{\alpha} - A_M((L_{TJ})_n)^{\alpha} > A_T((L_{TJ})_e)^{\alpha} - A_M((L_{TJ})_e)^{\alpha}. \tag{38}\]

The derivative of \(A_T((L_{TJ})_n)^{\alpha} - A_M((L_{TJ})_n)^{\alpha}\) with respect to \(L_{TJ}\) equals \(\alpha A_T(L_{TJ})^{\alpha-1} - A_M\), which decreases with \(L_{TJ}\). Thus, since \((L_{TJ})_n^* > (L_{TJ})_n\), the above condition holds if \(\alpha A_T((L_{TJ})_n)^{\alpha-1} - A_M \leq 0 \Leftrightarrow (L_{TJ})_n^* \geq \left( \frac{A_T}{A_M} \right)^{\frac{1}{\alpha-1}}\). Because \((L_{TJ})_n^*\) is the solution to (25), the condition holds if
\[
\frac{A_M}{\alpha} - \frac{1}{\beta \omega_s} \left[ 1 - \frac{2}{\theta} \left( \frac{A_T}{A_M} \right)^{\frac{1}{1-\alpha}} \right] \geq A_M. \tag{39}\]

Hence, \(Y_n^* > Y_e^*\) is true when \(\beta\) is not very large so that the inequality holds.

The derivative of the LHS of the above inequality with respect to \(\alpha\) equals
\[
- \frac{A_M}{\alpha^2} + \beta \omega_s \frac{2}{\theta} \left( \frac{1}{1-\alpha} \ln \left( \frac{A_T}{A_M} \right) \right) + \left( \frac{1}{1-\alpha} \ln \left( \frac{A_T}{A_M} \right) \right)^{\frac{1}{1-\alpha}} \tag{40}\]
and the second derivative equals
\[
2 \frac{A_M}{\alpha^3} + \beta \omega_s \frac{2}{\theta} \left( \frac{1}{1-\alpha} \ln \left( \frac{A_T}{A_M} \right) \right) + \left( \frac{1}{1-\alpha} \ln \left( \frac{A_T}{A_M} \right) \right)^{\frac{1}{1-\alpha}} \tag{41}\]
from \(\left( \frac{A_T}{A_M} \right)^{\frac{1}{1-\alpha}} > 1\).

Because (39) holds as \(\alpha \to 0\), does not hold as \(\alpha \to 1\), and the first derivative of the LHS of the
equation is $-\infty$ as $\alpha \to 0$, and it increases as $\alpha$ goes up, there exists a $\alpha \in (0, 1)$, which depends on exogenous variables and parameters, below which $(LT_J)_d^* \geq \left(\frac{\alpha \chi}{\lambda^s/t^s}\right)^{\frac{1}{1-\alpha}}$ and thus $Y^*_n > Y^*_e$ hold.

By contrast, $Y^*_n < Y^*_e$ holds when $(LT_J)_e^* \leq \left(\frac{\alpha \chi}{\lambda^s/t^s}\right)^{\frac{1}{1-\alpha}}$, which is always true if $(\frac{\alpha \chi}{\lambda^s/t^s})^{\frac{1}{1-\alpha}} \geq \frac{L}{n_e}$. Otherwise,

$$(LT_J)_e^* \leq \left(\frac{\alpha \chi}{\lambda^s/t^s}\right)^{\frac{1}{1-\alpha}} \Leftrightarrow \frac{\lambda^s}{\chi} - \beta \omega_1 \left[1 - 2n_e \left(\frac{\alpha \chi}{\lambda^s/t^s}\right)^{\frac{1}{1-\alpha}}\right] \leq \lambda M \text{ from (20).} \quad (42)$$

From the equation, when $\beta$ is large enough, $(LT_J)_e^* \leq \left(\frac{\alpha \chi}{\lambda^s/t^s}\right)^{\frac{1}{1-\alpha}}$ and thus $Y^*_n < Y^*_e$ hold. Also, from a similar reasoning as above, there exists a $\beta \in (0, 1)$ ($\beta > \alpha$) above which $Y^*_n < Y^*_e$ hold. ■

**Proof of Proposition 2.** (i) $F_n^* < F_{td}^*$ is from (24) and (30), $F_{md}^* < F_e^*$ is from (18) and (33), and $F_{td}^* < F_d^* < F_{md}^*$ is from (26), (30), and (33) and $(LT_J)_d^* < (LT_J)_d^*$ is shown in (ii).

(ii) $(LT_J)_d^* = (LT_J)_n^*$ and $(LT_J)_d^* = (LT_J)_e^*$ are shown in Sections 3.2.2 and 3.2.3. As shown in footnote 23, the LHS of (20), the indifference condition whose solution is $(LT_J)_e^*$, decreases with $LT_J$ for $LT_J \leq (LT_J)_e^*$.

Hence, $(LT_J)_d^* < (LT_J)_e^*$ holds, if the LHS of (27), the indifference condition whose solution is $(LT_J)_d^*$, is smaller than that of (20) at $LT_J = (LT_J)_d^*$, which is true because

$$\beta \left[\left(\gamma_0 + \eta_1 F_d(LT_J)\right)\frac{n_e - 1}{n_e} + \omega_1 \left(n_e \left(\frac{LT_J}{L}\right)^2 - (1 - n_e \left(\frac{LT_J}{L}\right))\right)^2\right]$$

$$- \frac{1}{2} \left(\frac{n_e - 1}{n_e}\right)^{\frac{1}{\theta-1}} \left[\left(\delta F_d(LT_J)\right)^{\frac{1}{\theta-1}} - \left(\delta F_d(LT_J) - \beta \eta_1\right)^{\frac{1}{\theta-1}}\right] - \gamma \Delta S \leq - \beta \omega_1 (1 - 2n_e \left(\frac{LT_J}{L}\right)) \quad (43)$$

$$\Leftrightarrow \beta \left[\left(\gamma_0 + \eta_1 F_d(LT_J)\right)\frac{n_e - 1}{n_e} - \omega_1 n_e (n_e - 1) \left(\frac{LT_J}{L}\right)^2\right] - \frac{1}{2} \left(\frac{n_e - 1}{n_e}\right)^{\frac{1}{\theta-1}} \left[\left(\delta F_d(LT_J)\right)^{\frac{1}{\theta-1}} - \left(\delta F_d(LT_J) - \beta \eta_1\right)^{\frac{1}{\theta-1}}\right] \leq \gamma \Delta S, \quad (44)$$

where the inequality holds from (77) in the proof of Proposition A2.

As shown in footnote 24, the shape of the LHS of (25), the indifference condition whose solution is $(LT_J)_n^*$, is similar to that of (20). Hence, $(LT_J)_d^* > (LT_J)_n^*$ holds if the LHS of (27) is greater than that of (25) at $LT_J = (LT_J)_d^*$, which is true because

$$\beta \left[\left(\gamma_0 + \eta_1 F_d(LT_J)\right)\frac{n_e - 1}{n_e} + \omega_1 \left(n_e \left(\frac{LT_J}{L}\right)^2 - (1 - n_e \left(\frac{LT_J}{L}\right))\right)^2\right]$$

$$- \frac{1}{2} \left(\frac{n_e - 1}{n_e}\right)^{\frac{1}{\theta-1}} \left[\left(\delta F_d(LT_J)\right)^{\frac{1}{\theta-1}} - \left(\delta F_d(LT_J) - \beta \eta_1\right)^{\frac{1}{\theta-1}}\right] - \gamma \Delta S \geq - \beta \omega_1 (1 - 2n_e \left(\frac{LT_J}{L}\right)) \quad (45)$$

$$\Leftrightarrow \beta \left[\left(\gamma_0 + \eta_1 F_d(LT_J)\right)\frac{n_e - 1}{n_e} - \omega_1 n_e (n_e - 1) \left(\frac{LT_J}{L}\right)^2\right] - \frac{1}{2} \left(\frac{n_e - 1}{n_e}\right)^{\frac{1}{\theta-1}} \left[\left(\delta F_d(LT_J)\right)^{\frac{1}{\theta-1}} - \left(\delta F_d(LT_J) - \beta \eta_1\right)^{\frac{1}{\theta-1}}\right] \geq \gamma \Delta S, \quad (46)$$

where the inequality holds from (78) in the proof of Proposition A2.

(iii) Denote the total output of private goods in equilibrium (d) by $Y_d^*$. From (1) and (2), $Y_n^* > Y_d^* > Y_e^* \Leftrightarrow AT((LT_J)_n^* \alpha - AM(LT_J)_n) > AT((LT_J)_d^* \alpha - AM(LT_J)_d) > AT((LT_J)_e^* \alpha - AM(LT_J)_e)$. Because $(LT_J)_n^* < (LT_J)_d^* < (LT_J)_e^*$, if $\beta$ is not very large so that (39) in the proof of Proposition 1 (iii) holds, $AT((LT_J)_n^* \alpha - AM(LT_J))$ decreases with $LT_J$ for $LT_J \geq (LT_J)_n^*$ and thus $Y_n^* > Y_d^* > Y_e^*$ is true.
If $\beta$ is large enough that (42) in the proof of the proposition holds, $A_T(L_{TJ})^\alpha - A_M L_{TJ}$ increases with $L_{TJ}$ for $L_{TJ} \leq (L_{TJ})^*_e$ and thus $Y^*_n < Y^*_d < Y^*_e$ is true. As for the relationship between $\alpha$ and the magnitude relation of $Y^*_n$ and $Y^*_d$ or $Y^*_e$, the corresponding proof of Proposition 1 applies since $(L_{TJ})^*_n > (L_{TJ})^*_d > (L_{TJ})^*_e$ holds. $Y^*_n = Y^*_a$ and $Y^*_d = Y^*_e$ are evident from $(L_{TJ})^*_n = (L_{TJ})^*_a$ and $(L_{TJ})^*_d = (L_{TJ})^*_e$.

(iv) The total cost of conflict is $(L_n$ and $L_e$ are respectively numbers of those identifying with the nation and their ethnic group) $\frac{1}{\theta}[(f_{i,n})^\theta L_n + (f_{i,e})^\theta L_e] = \frac{1}{\theta}[(f_{i,n})^\theta - f_{i,n} L_n + (f_{i,e})^\theta - f_{i,e} L_e] = \frac{1}{\theta} \n_e \left[ (\delta V_{f_T} - \beta \eta_1) f_{i,n} L_n + \delta V_{f_{TJ}} f_{i,e} L_e \right]$, where the second equality is from the first order conditions of utility maximization and $F_{e} = \frac{\n_e - 1}{\theta} F$. The total cost decreases with $L_n$ since $f_{i,n}$ increases with $L_n$ from (22) and Proposition 2. From this result and (iii), aggregate material payoff increases with $L_n$ unless $\alpha$ or $\beta$ is very high.

**Proof of Proposition 4.** It is enough to prove that the terms on the opposite side of $\gamma \Delta S$ of all equilibrium conditions— the LHSs of (62) and (73) in the proof of Proposition A1 and of (77) and (78) in the proof of Proposition A2— increase with $\eta_0$ (that is, all dividing lines in Figure 4 shift upward with an increase in $\eta_0$). As for the homogeneous identity equilibria, since $F$ is independent of $\eta_0$, the result is straightforward from (62) and (73). As for equilibrium (d), since $F$ is independent of $\eta_0$ for given $L_{TJ}$ from (26), the result is straightforward from (77) and (78).

(Each term of the conditions of the remaining equilibria are same as one of these terms.)

**Proof of Proposition 5.** (i) Straightforward from the equation determining $F$ of each equilibrium, (18), (24), (26), (27), (29), and (32).

(ii) It is enough to prove that the terms on the opposite side of $\gamma \Delta S$ of all equilibrium conditions—the LHSs of (62) and (73) in the proof of Proposition A1 and of (77) and (78) in the proof of Proposition A2— increase with $V$.

[Equilibrium (e)] The derivative of the LHS of (62) in the proof of Proposition A1 with respect to $V$ is, from (18),

$$
\frac{1}{\theta} (V)^{-1} \left\{ \beta \eta_1 \n_e \n_e \left( \frac{V}{F_e} \right)^{\frac{\theta - 1}{\theta}} \left[ \delta V_{f_T}^{\frac{\theta - 1}{\theta}} - \delta V_{f_{TJ}}^{\frac{\theta - 1}{\theta}} \right] \right\} \delta V
$$

$$
> \frac{1}{\theta} (V)^{-1} \left\{ \beta \eta n_e \n_e \left( \frac{V}{F_e} \right)^{\frac{\theta - 1}{\theta}} \right\} = \frac{1}{\theta} (V)^{-1} \left( \frac{V}{F_e} \right)^{\frac{\theta - 1}{\theta}} \delta V_{f_{TJ}}^{\frac{\theta - 1}{\theta}}. \quad (47)
$$

For $F_e = \left( \frac{\delta V_{f_T}}{\theta} \right)^{\frac{\theta - 1}{\theta}} L$ and $F_n = \left( \frac{\delta V_{f_{TJ}} - \beta \eta_1}{\theta} \right)^{\frac{\theta - 1}{\theta}} L$ not to be too close, $\beta \eta_1$ must be of a similar order of magnitude to $\delta V_{f_T}$ and $\delta V_{f_{TJ}}$. Then, $\beta \eta_1 - \frac{1}{L} \delta V_{f_{TJ}} > 0$ and the derivative is positive.

[Equilibrium (n)] Since $\frac{d F_n}{d V} = \frac{1}{\theta} \frac{V}{\theta - 1} \delta V_{f_{TJ}}^{\frac{\theta - 1}{\theta}} - \beta \eta_1$ and $\frac{d (V_n)}{d V} = \frac{1}{\theta} \frac{V}{\theta - 1} \delta V_{f_{TJ}}^{\frac{\theta - 1}{\theta}} - \beta \eta_1$ from (24), the derivative of the LHS of (73) with respect to $V$ is,
Proof of Proposition A1.

(i) Equilibrium (e): [Sector M] The utility of individual i of ethnic group J in sector M equals, from (15) and (16),

\[
A_M - \frac{1}{\theta} (f_{i,e})^\theta + \delta \frac{F_J}{\theta} V - \beta \omega_s \left( \frac{L_{TJ}}{L/n_e} \right)^2 + \gamma S_E.
\]

If he deviates and identifies with the nation, the highest utility he gets is

\[
A_M - \frac{1}{\theta} (f_{i,n})^\theta + \delta \frac{(F_J)^'}{\theta} V - \beta \left( \eta_0 + \eta_1 F \right) \frac{n_e - 1}{n_e} + \omega_s \left[ \left( \frac{L_{TJ}}{L} \right)^2 + \sum_{K \neq J} \left( \frac{L_{TK}}{L} \right)^2 \right] + \gamma S_N,
\]

(Each term of the equilibrium condition of the remaining equilibria are same as one of the above terms.)

(ii) Equilibrium (f): From (26),

\[
\frac{dF_d(L_{TJ})}{dV} = \frac{1}{\theta - 1} \frac{\delta V}{F_d(L_{TJ})} \left[ \left( \delta F_d(L_{TJ}) \right)^{\frac{1}{\theta - 1}} n_e L_{TJ} + \left( \delta F_d(L_{TJ}) - \beta \eta_1 \right)^{\frac{1}{\theta - 1}} (L - n_e L_{TJ}) \right]
\]

(49)

Thus, the derivative of the LHS of (77) or (78) with respect to V is,

\[
\frac{1}{\theta - 1} \frac{\delta V}{F_d(L_{TJ})} \left[ \left( \delta F_d(L_{TJ}) \right)^{\frac{1}{\theta - 1}} n_e L_{TJ} + \left( \delta F_d(L_{TJ}) - \beta \eta_1 \right)^{\frac{1}{\theta - 1}} (L - n_e L_{TJ}) \right]
\]

(50)

\[
1 + \frac{1}{\theta - 1} \frac{\delta V}{F_d(L_{TJ})} \left[ \left( \delta F_d(L_{TJ}) \right)^{\frac{1}{\theta - 1}} n_e L_{TJ} + \left( \delta F_d(L_{TJ}) - \beta \eta_1 \right)^{\frac{1}{\theta - 1}} (L - n_e L_{TJ}) \right]
\]

\[
\times \left[ \beta \eta_1 - \frac{\left( \delta F_d(L_{TJ}) \right)^{\frac{1}{\theta - 1}} n_e L_{TJ} + \left( \delta F_d(L_{TJ}) - \beta \eta_1 \right)^{\frac{1}{\theta - 1}} (L - n_e L_{TJ})}{\left( \delta F_d(L_{TJ}) \right)^{\frac{1}{\theta - 1}} n_e L_{TJ} + \left( \delta F_d(L_{TJ}) - \beta \eta_1 \right)^{\frac{1}{\theta - 1}} (L - n_e L_{TJ})} \right],
\]

(51)

where the expression inside the large square bracket is greater than

\[
\beta \eta_1 - \frac{\left( \delta F_d(L_{TJ}) \right)^{\frac{1}{\theta - 1}} n_e L_{TJ} + \left( \delta F_d(L_{TJ}) - \beta \eta_1 \right)^{\frac{1}{\theta - 1}} (L - n_e L_{TJ})}{\left( \delta F_d(L_{TJ}) \right)^{\frac{1}{\theta - 1}} n_e L_{TJ} + \left( \delta F_d(L_{TJ}) - \beta \eta_1 \right)^{\frac{1}{\theta - 1}} (L - n_e L_{TJ})} > \beta \eta_1 - \frac{1}{\theta} \delta \frac{V}{F_d(L_{TJ})} > 0.
\]

(52)

(Each term of the equilibrium condition of the remaining equilibria are same as one of the above terms.)
where \((f_{i,n})' = \left[\frac{F_j}{(F_j')^2}V - \beta \eta_1 n_e^{-1}\right]^{\frac{1}{\beta}}\), not \((f_{i,n})' = 0\), from the assumption (14), \((F_j)' = (f_{i,n})' + \left(\frac{L}{n_e} - 1\right)f_{i,e}\), and \(F' = (F_j)' + F_{-j}\).

When \(L\) is large enough, the deviation by one player affects aggregate values \((F_j)'\) and \(F'\) very little, thus the above equation is approximated very well by the following equation that is marginally larger than the original one

\[
AM - \frac{1}{\theta}(f_{i,n})^\theta + \delta \frac{F_j}{F}V - \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_{e-1}}{n_e} + \omega_s \left[ \left( \frac{L_{TJ}}{L} \right)^2 + \sum_{K \neq j} \left( \frac{L_{TK}}{L} \right)^2 \right] \right\} + \gamma S_N, \text{ where } f_{i,n} \text{ is given by (22).}
\]

Thus, the deviation is not profitable if

\[
-\frac{1}{\theta}(f_{i,e})^\theta - \beta \omega_s \left( \frac{L_{TJ}}{L/n_e} \right)^2 + \gamma S_E \geq -\frac{1}{\theta}(f_{i,n})^\theta - \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_{e-1}}{n_e} + \omega_s \left[ \left( \frac{L_{TJ}}{L} \right)^2 + \sum_{K \neq j} \left( \frac{L_{TK}}{L} \right)^2 \right] \right\} + \gamma S_N
\]

\[\Leftrightarrow \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_{e-1}}{n_e} + \omega_s \left[ \sum_{K \neq j} \left( \frac{L_{TK}}{L} \right)^2 - (n_e - 1) \left( \frac{L_{TJ}}{L} \right)^2 \right] \right\} - \frac{1}{\theta}(f_{i,e})^\theta - (f_{i,n})^\theta \geq \gamma \Delta S.\]

\[\text{[Sector TJ]} \text{ The utility of individual } i \text{ of ethnic group } J \text{ in sector } TJ \text{ is, from (17) and (16),}
\]

\[A_T (L_{TJ})^{n-1} - \frac{1}{\theta}(f_{i,n})^\theta + \delta \frac{F_j}{F}V - \beta \omega_s \left( 1 - \frac{L_{TJ}}{L/n_e} \right)^2 + \gamma S_E\]

If he deviates and identifies with the nation, the highest utility is well approximated by

\[A_T (L_{TJ})^{n-1} - \frac{1}{\theta}(f_{i,n})^\theta + \delta \frac{F_j}{F}V - \beta \omega_s \left[ 1 - \frac{L_{TJ}}{L} \right]^2 + \sum_{K \neq j} \left( \frac{L_{TK}}{L} \right)^2 \right] + \gamma S_N.
\]

The deviation is not profitable if

\[
-\frac{1}{\theta}(f_{i,e})^\theta - \beta \omega_s \left[ 1 - \frac{L_{TJ}}{L/n_e} \right]^2 + \gamma S_E \geq -\frac{1}{\theta}(f_{i,n})^\theta - \beta \left\{ \omega_s \frac{n_{e-1}}{n_e} + \omega_s \left[ 1 - \frac{L_{TJ}}{L} \right]^2 + \sum_{K \neq j} \left( \frac{L_{TK}}{L} \right)^2 \right\} + \gamma S_N
\]

\[\Leftrightarrow \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_{e-1}}{n_e} + \omega_s \left[ \sum_{K \neq j} \left( \frac{L_{TK}}{L} \right)^2 - (n_e - 1) \left( \frac{L_{TJ}}{L} \right)^2 \right] \right\} - \frac{1}{\theta}(f_{i,e})^\theta - (f_{i,n})^\theta \geq \gamma \Delta S.\]

\[\text{[The equilibrium condition]} \text{ From this equation and (57), if the condition for the modern sector holds, so does the one for the traditional sector. Hence, (57) is the condition for the existence of the equilibrium when } L_{TJ} \text{ is the solution for (20). Since ethnic groups are symmetric and thus values of aggregate variables of all groups are the same, (57) becomes}
\]

\[\beta \left[ (\eta_0 + \eta_1 F_e) \frac{n_{e-1}}{n_e} - \omega_s n_e (n_e - 1) \left( \frac{L_{TJ}}{L} \right)^2 \right] - \frac{1}{\theta} \left[ (n_e - 1) \left( \frac{L_{TJ}}{L} \right)^2 \right] \geq \gamma \Delta S
\]

\[\Leftrightarrow \beta \Delta d^2 \left[ F_e^* - n_e (n_e - 1) \left( \frac{L_{TJ}}{L} \right)^2 \right] - \Delta c(F_e^*) \geq \gamma \Delta S,
\]

where \(F_e^*\) is given by (18), and \(\Delta d^2[\cdot], \Delta c(\cdot), \text{ and } \Delta S \text{ are as defined just before the proposition.}

The condition holds for any \(L_{TJ} \in [0, \frac{L}{n_e}]\) when \(\gamma \Delta S \leq \beta \Delta d^2 \left[ F_e^* - \frac{n_{e-1}}{n_e} \right] - \Delta c(F_e^*)\). When \(\gamma \Delta S \in\)
\[
(\beta \Delta d^2 [F^e_* - \frac{n_e - 1}{n_e}] - \Delta c(F^e_*), \beta \Delta d^2 [F^e_* - 0] - \Delta c(F^e_*)], \text{ the condition holds for } L_{TJ} \in [0, (L_{TJ})^*], \text{ where}
\]

\[(L_{TJ})^* \text{ is the solution for } \beta \Delta d^2 [F^e_* - n_e(n_e - 1)\left(\frac{(f_{i,e})''}{L}\right)^2] - \Delta c(F^e_*) = \gamma \Delta S.\]

(ii) Equilibrium (n): [Sector TJ] The utility of individual i of ethnic group J in sector TJ is, from (23) and (22),

\[
A_T(L_{TJ})^{\alpha - 1} - \frac{1}{\theta} (f_{i,n})^\theta + \delta F_j F^n V - \beta \left(\omega_e - \frac{n_e - 1}{n_e} + \omega_s \left\{1 - \frac{L_{TJ}}{L}\right\} + \sum_{K \neq J} \left(\frac{L_{TK}}{L}\right)^2\right) + \gamma S_N. \quad (64)
\]

If he deviates and identifies with his or her ethnic group, the highest utility he gets is

\[
A_T(L_{TJ})^{\alpha - 1} - \frac{1}{\theta} (f_{i,e})^\theta + \delta F_j F^n V - \beta \omega_s \left(1 - \frac{L_{TJ}}{L/n_e}\right)^2 + \gamma S_E, \quad (65)
\]

where \((f_{i,e})'' = \left[\frac{\delta F_j F^n V}{(F^n)^2}\right]^\frac{1}{n_e} - 1, (F_j)' = (f_{i,e})' + \frac{L}{n_e} - 1 f_{i,n}, \text{ and } (F)' = (f_{i,e})' + F_{-J}.

When \(L\) is large enough, the deviation by one player affects aggregate values \((F_j)'\) and \((F)'\) very little, thus the above equation is approximated very well by the following equation that is marginally smaller than the original one

\[
A_T(L_{TJ})^{\alpha - 1} - \frac{1}{\theta} (f_{i,e})^\theta + \delta F_j F^n V - \beta \omega_s \left(1 - \frac{L_{TJ}}{L/n_e}\right)^2 + \gamma S_E, \quad (66)
\]

The deviation is not profitable if\(^{58}\)

\[
- \frac{1}{\theta} (f_{i,n})^\theta - \beta \left(\omega_e - \frac{n_e - 1}{n_e} + \omega_s \left\{1 - \frac{L_{TJ}}{L}\right\} + \sum_{K \neq J} \left(\frac{L_{TK}}{L}\right)^2\right) + \gamma S_N > - \frac{1}{\theta} (f_{i,e})^\theta - \beta \omega_s \left(1 - \frac{L_{TJ}}{L/n_e}\right)^2 + \gamma S_E \quad (67)
\]

\[
\iff \beta \left(\eta_0 + \eta_1 F\right)n_e - 1 + \omega_s \left\{\sum_{K \neq J} \left(\frac{L_{TK}}{L}\right)^2 + \left(n_e - 1\right)\frac{L_{TJ}}{L}\left[2 - \left(n_e - 1\right)\frac{L_{TJ}}{L}\right]\right\} - \frac{1}{\theta} [\left(f_{i,e}\right)^\theta - (f_{i,n})^\theta] < \gamma \Delta S. \quad (68)
\]

[Sector M] The utility of individual i of ethnic group J in sector M equals, from (21) and (22),

\[
A_M - \frac{1}{\theta} (f_{i,n})^\theta + \delta F_j F^n V - \beta \left(\eta_0 + \eta_1 F\right)n_e - 1 + \omega_s \left[\left(\frac{L_{TJ}}{L}\right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{L}\right)^2\right] + \gamma S_N. \quad (69)
\]

If he deviates and identifies with his or her ethnic group, the highest utility is well approximated by

\[
A_M - \frac{1}{\theta} (f_{i,e})^\theta + \delta F_j F^n V - \beta \omega_s \left(\frac{L_{TJ}}{L/n_e}\right)^2 + \gamma S_E. \quad (70)
\]

The deviation is not profitable if

\[
- \frac{1}{\theta} (f_{i,n})^\theta - \beta \left(\eta_0 + \eta_1 F\right)n_e - 1 + \omega_s \left[\sum_{K \neq J} \left(\frac{L_{TK}}{L}\right)^2 - \left(n_e - 1\right)\left(\frac{L_{TJ}}{L}\right)^2\right] + \gamma S_N > - \frac{1}{\theta} (f_{i,e})^\theta - \beta \omega_s \left(\frac{L_{TJ}}{L/n_e}\right)^2 + \gamma S_E \quad (71)
\]

\[
\iff \beta \left(\eta_0 + \eta_1 F\right)n_e - 1 + \omega_s \left\{\sum_{K \neq J} \left(\frac{L_{TK}}{L}\right)^2 - \left(n_e - 1\right)\left(\frac{L_{TJ}}{L}\right)^2\right\} - \frac{1}{\theta} [\left(f_{i,e}\right)^\theta - (f_{i,n})^\theta] < \gamma \Delta S. \quad (72)
\]

\(^{58}\) The equation must hold with strict inequality because the deviant’s approximate utility is marginally smaller than the true utility.
[The equilibrium condition] From this equation and (68), if the condition for sector $TJ$ holds, so does the one for sector $M$. Hence, (68) is the condition for the existence of this equilibrium when $L_{TJ}$ is the solution for (25). Since ethnic groups are symmetric, (68) becomes

$$
\beta \left[ (\eta_0 + \eta_1 F_n^*) \frac{n_e - 1}{n_e} + \omega(n_e - 1) \frac{L_{TJ}^2}{L} - \frac{1}{2} \left( \frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \left[ \left( \delta \frac{V}{F_n^*} \right)^{\frac{\theta}{\theta - 1}} - \left( \delta \frac{V}{F_n^*} - \beta \eta_1 \right)^{\frac{\theta}{\theta - 1}} \right] \right] < \gamma \Delta S \quad (73)
$$

$$\Leftrightarrow \beta \Delta d^2 \left[ F_n^*, (n_e - 1) \frac{L_{TJ}^2}{L} - \frac{L_{TJ}}{L} \right] - \Delta c(F_n^*) < \gamma \Delta S, \quad (74)
$$

where $F_n^*$ is the solution for (24).

The above inequality holds for any $L_{TJ} \in [0, \frac{L}{n_e}]$ when $\gamma \Delta S > \beta \Delta d^2 \left[ F_n^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_n^*)$, and for $L_{TJ} \in [0, (L_{TJ})^{\text{opt}}]$ when $\gamma \Delta S \in \left( \beta \Delta d^2 \left[ F_n^*, 0 \right] - \Delta c(F_n^*) \right) \beta \Delta d^2 \left[ F_n^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_n^*)$, where $(L_{TJ})^{\text{opt}}$ is the solution for $\beta \Delta d^2 \left[ F_n^*, (n_e - 1) \frac{L_{TJ}^2}{L} - \frac{L_{TJ}}{L} \right] - \Delta c(F_n^*) = \gamma \Delta S$. 

**Proof of Proposition A2.** [The proof that no other heterogeneous identity equilibria exist] If workers in sector $M$ weakly prefer to identify with their ethnic group, from (57), the following must hold in a symmetric equilibrium:

$$
\beta \left[ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} - \omega(n_e - 1) \left( \frac{L_{TJ}^2}{L} \right)^2 \right] - \frac{1}{2} \left( \frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \left[ \frac{V}{F} \right]^{\frac{\theta}{\theta - 1}} - \left( \frac{V}{F} - \beta \eta \right)^{\frac{\theta}{\theta - 1}} \right] \geq \gamma \Delta S. \quad (75)
$$

If workers in sector $TJ$ weakly prefer to identify with the nation, from (68), the following must hold in a symmetric equilibrium:

$$
\beta \left[ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega(n_e - 1) \frac{L_{TJ}^2}{L} (2 - n_e \frac{L_{TJ}}{L}) \right] - \frac{1}{2} \left( \frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \left[ \frac{V}{F} \right]^{\frac{\theta}{\theta - 1}} - \left( \frac{V}{F} - \beta \eta \right)^{\frac{\theta}{\theta - 1}} \right] \leq \gamma \Delta S. \quad (76)
$$

Both conditions cannot hold simultaneously and thus such situations do not arise in equilibrium.

(i) Equilibrium (d): [Sector $M$] Because sector $M$ workers identify with the nation, the condition for them not to deviate from the equilibrium is given by (72) as in equilibrium (n). In the symmetric equilibrium, the equation becomes

$$
\beta \left[ (\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e - 1}{n_e} - \omega(n_e - 1)n_e \left( \frac{L_{TJ}^2}{L} \right) \right] - \frac{1}{2} \left( \frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \left[ \frac{V}{F_d(L_{TJ})} \right]^{\frac{\theta}{\theta - 1}} - \left( \frac{V}{F_d(L_{TJ})} - \beta \eta_1 \right)^{\frac{\theta}{\theta - 1}} \right] < \gamma \Delta S, \quad (77)
$$

$$\Leftrightarrow \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1)n_e \left( \frac{L_{TJ}^2}{L} \right) \right] - \Delta c(F_d(L_{TJ})) < \gamma \Delta S,
$$

where $F_d(L_{TJ})$ is the solution for (26) and increases with $L_{TJ}$.

The relation between the LHS of the equation and $L_{TJ}$ is ambiguous, but the relation is positive for small $L_{TJ}$ because the derivative of the LHS at $L_{TJ} = 0$ is positive.

[Sector $TJ$] Because sector $TJ$ workers identify with their ethnic group, the condition for them not to deviate from the equilibrium is given by (61) as in equilibrium (e). In the symmetric equilibrium, the equation becomes
\[
\beta \left[ \left( \eta_0 + \eta_1 F_d(L_{TJ}) \right) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{L} \right] - \frac{1}{g} \left( \frac{n_e - 1}{n_e} \right)^{\frac{\alpha}{\theta}} \left[ \left( \frac{1}{F_d(L_{TJ})} \right)^{\frac{\alpha}{\theta}} - \left( \frac{1}{F_d(V)} - \beta \eta \right)^{\frac{\alpha}{\theta}} \right] \geq \gamma S_N,
\]

\[
(78)
\]

\[
\Leftrightarrow \beta \Delta d^2 \left[ F_d(L_{TJ}), (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \Delta c(F_d(L_{TJ})) \geq \gamma S_N,
\]

\[
(79)
\]

where the LHS increases with \( L_{TJ} \) from (26).

Thus, the condition holds for any \( L_{TJ} \in \left[ 0, \frac{L}{n_e} \right] \) when \( \gamma S_N \leq \beta \Delta d^2[F_n^*, 0] - \Delta c(F_n^*) \), and for \( L_{TJ} \in \left[ (L_{TJ})^0, \frac{L}{n_e} \right] \) when \( \gamma S_N \leq \beta \Delta d^2[F_n^*, 0] - \Delta c(F_n^*) \), \( \beta \Delta d^2 \left[ F_n^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_n^*) \), where \( (L_{TJ})^0 \) is the solution for \( \beta \Delta d^2 \left[ F_d(L_{TJ}), (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \Delta c(F_d(L_{TJ})) = \gamma S_N \).

The equilibrium condition hence, the equilibrium exists iff \( \beta \Delta d^2 \left[ F_d(L_{TJ}), (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \Delta c(F_d(L_{TJ})) < \gamma S_N \leq \beta \Delta d^2 \left[ F_d(L_{TJ}), (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \Delta c(F_d(L_{TJ})) \), where \( F_d(L_{TJ}) \) is the solution for (26) and increases with \( L_{TJ} \) and \( L_{TJ} = (L_{TJ})^0 \) is the solution for (27).

To be more detailed, the equilibrium exists for \( L_{TJ} \in \left[ (L_{TJ})^0, \frac{L}{n_e} \right] \) when

\[
\gamma S < \beta \Delta d^2 \left[ F_d(L_{TJ}), (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \Delta c(F_d(L_{TJ})) < \gamma S \text{ when } \gamma S \in \left( \min_{L_{TJ}} \left\{ \beta \Delta d^2 \left[ F_d(L_{TJ}), (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \Delta c(F_d(L_{TJ})) \right\}, \max_{L_{TJ}} \left\{ \beta \Delta d^2 \left[ F_d(L_{TJ}), (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \Delta c(F_d(L_{TJ})) \right\} \right).
\]

(ii) Equilibrium (Md): [Sector M] As shown in Section 3.2.2, the following indifference condition for identity choices of sector M workers must hold

\[
\beta \left[ \left( \eta_0 + \eta_1 F \right) \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \frac{L_{TJ}}{L} \right] - \frac{1}{g} \left( \frac{n_e - 1}{n_e} \right)^{\frac{\alpha}{\theta}} \left[ \left( \frac{1}{F} \right)^{\frac{\alpha}{\theta}} - \left( \frac{1}{F_d(V)} - \beta \eta \right)^{\frac{\alpha}{\theta}} \right] = \gamma S,
\]

where \( F \) satisfies

\[
F = \left( \frac{n_e - 1}{n_e} \right)^{\frac{1}{\theta}} \left\{ \left( \frac{1}{F} - \beta \eta \right)^{\frac{1}{\theta}} P_{M,n} \left( L - n_e L_{TJ} \right) + \left( \frac{1}{F_d(V)} - \beta \eta \right)^{\frac{1}{\theta}} \left( n_e L_{TJ} + (1 - P_{M,n}) (L - n_e L_{TJ}) \right) \right\}.
\]

(30)

Given \( L_{TJ} \), the LHS of (29) increases with \( F \) and \( F \) satisfying (30) decreases with \( P_{M,n} \). Hence, \( F \) and \( P_{M,n} \) satisfying both equations exist iff

\[
\beta \left[ \left( \eta_0 + \eta_1 F_d(L_{TJ}) \right) \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \frac{L_{TJ}}{L} \right] - \frac{1}{g} \left( \frac{n_e - 1}{n_e} \right)^{\frac{1}{\theta}} \left[ \left( \frac{1}{F_d(L_{TJ})} \right)^{\frac{1}{\theta}} - \left( \frac{1}{F_d(V)} - \beta \eta \right)^{\frac{1}{\theta}} \right] < \gamma S < \beta \left[ \left( \eta_0 + \eta_1 F_d(L_{TJ}) \right) \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \frac{L_{TJ}}{L} \right] - \frac{1}{g} \left( \frac{n_e - 1}{n_e} \right)^{\frac{1}{\theta}} \left[ \left( \frac{1}{F_d(L_{TJ})} \right)^{\frac{1}{\theta}} - \left( \frac{1}{F_d(V)} - \beta \eta \right)^{\frac{1}{\theta}} \right],
\]

(80)

where \( F^*_e \) is given by (18) and \( F_d(L_{TJ}) \) is given by (26) and increases with \( L_{TJ} \).

The second inequality of (80) holds for any \( L_{TJ} \in \left[ 0, \frac{L}{n_e} \right] \) when \( \gamma S < \beta \Delta d^2 \left[ F_e^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*) \), and for \( L_{TJ} \in \left[ \left( L_{TJ} \right)^{\dagger}, \frac{L}{n_e} \right] \) when \( \gamma S \in \left[ \beta \Delta d^2 \left[ F_e^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*) \right], \beta \Delta d^2 \left[ F_e^*, 0 \right] - \Delta c(F_e^*) \), where \( \left( L_{TJ} \right)^{\dagger} \) is \( L_{TJ} \) satisfying \( \beta \Delta d^2 \left[ F_e^*, (n_e - 1) \frac{L_{TJ}}{L} \right] - \Delta c(F_e^*) = \gamma S \).

The LHS of the first inequality is same as (27) in (i), thus the relation between the LHS and \( L_{TJ} \) is positive for small \( L_{TJ} \) but generally ambiguous.
[Sector TJ] Because workers in sector TJ identify with their ethnic group, the condition for them not to deviate from the equilibrium is given by (61) as in equilibrium (e). In the symmetric equilibrium, the condition becomes

$$\beta \left[ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \frac{1}{\beta} \left[ (\delta \frac{n_e - 1}{n_e} V) \frac{\theta}{\gamma} - (\delta \frac{n_e - 1}{n_e} V - \beta \eta_1 \frac{n_e - 1}{n_e}) \frac{\theta}{\gamma} \right] \geq \gamma \Delta S,$$  

where $F_T$ is the solution for (29) and (20). When (80) and thus (29) hold, it holds clearly.

The equilibrium condition Hence, from (80), the equilibrium exists iff $\beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1) n_e \left( \frac{L_{TJ}}{L} \right)^2 \right] - \Delta c(F_d(L_{TJ})) < \gamma \Delta S < \beta \Delta d^2 \left[ F_e^*, -(n_e - 1) \left( \frac{L_{TJ}}{L} \right)^2 \right] - \Delta c(F_e^*)$, where $F_d(L_{TJ})$ is the solution for (26) and increases with $L_{TJ}$ and $L_{TJ} = (L_{TJ})^*$ is the solution for (20).

To be more detailed, the equilibrium exists for $L_{TJ} \in [0, (L_{TJ})^{\dagger*}]$ when

$$\gamma \Delta S \in \left[ \max_{L_{TJ}} \left\{ \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1) n_e \left( \frac{L_{TJ}}{L} \right)^2 \right] - \Delta c(F_d(L_{TJ})) \right\}, \beta \Delta d^2 \left[ F_e^*, -(n_e - 1) n_e \left( \frac{L_{TJ}}{L} \right)^2 \right] - \Delta c(F_e^*) \right]$$

and for $L_{TJ} < \min \{ (L_{TJ})^{\dagger*}, \frac{L}{n_e} \}$ satisfying $\beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1) n_e \left( \frac{L_{TJ}}{L} \right)^2 \right] - \Delta c(F_d(L_{TJ})) < \gamma \Delta S$ when $\gamma \Delta S \in \left( \min_{L_{TJ}} \left\{ \beta \Delta d^2 \left[ F_e^*, 0 \right] - \Delta c(F_e^*), \beta \Delta d^2 \left[ F_e^*, -(n_e - 1) \right] - \Delta c(F_e^*) \right\}, \max_{L_{TJ}} \left\{ \beta \Delta d^2 \left[ F_d(L_{TJ}), -(n_e - 1) n_e \left( \frac{L_{TJ}}{L} \right)^2 \right] - \Delta c(F_d(L_{TJ})) \right\} \right)$,

where $(L_{TJ})^{\dagger*}$ is $L_{TJ}$ satisfying $\beta \Delta d^2 \left[ F_e^*, -(n_e - 1) \left( \frac{L_{TJ}}{L} \right)^2 \right] - \Delta c(F_e^*) = \gamma \Delta S$.

(iii) Equilibrium (Td): [Sector TJ] From Section 3.2.3, the indifference condition for identity choices must hold:

$$\beta \left[ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \frac{1}{\beta} \left( \frac{n_e - 1}{n_e} \right)^{\theta - 1} \left[ (\delta \frac{V}{TJ} \frac{\theta}{\gamma} - (\delta \frac{V}{TJ} - \beta \eta_1) \frac{\theta}{\gamma} \right] = \gamma \Delta S,$$  

where $F$ satisfies

$$F = \left( \frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\gamma - 1}} \left[ (\delta \frac{V}{TJ} - \beta \eta_1) \frac{1}{\gamma - 1} \left[ P_{TJ,n} n_e L_{TJ} + (L - n_e L_{TJ}) \right] + (\delta \frac{V}{TJ} \frac{\theta}{\gamma - 1} (1 - P_{TJ,n}) n_e L_{TJ} \right].$$  

Given $L_{TJ}$, the LHS of (32) increases with $F_T$ and $F_T$ satisfying (33) decreases with $P_{TJ,n}$. Hence, $F_T$ and $P_{TJ,n}$ satisfying both equations exist iff

$$\beta \left[ (\eta_0 + \eta_1 F_T) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \frac{1}{\beta} \left( \frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\gamma - 1}} \left[ (\delta \frac{V}{TJ} \frac{\theta}{\gamma} - (\delta \frac{V}{TJ} - \beta \eta_1) \frac{\theta}{\gamma} \right] < \gamma \Delta S$$

$$< \beta \left[ (\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{L} \left( 2 - n_e \frac{L_{TJ}}{L} \right) \right] - \frac{1}{\beta} \left( \frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\gamma - 1}} \left[ (\delta \frac{V}{TJ} \frac{\theta}{\gamma} - (\delta \frac{V}{TJ} - \beta \eta_1) \frac{\theta}{\gamma} \right],$$

where $F_T$ is given by (24) and $F_d(L_{TJ})$ is given by (26) and increases with $L_{TJ}$.

Thus, the first inequality of (82) holds for any $L_{TJ} \in [0, \frac{L}{n_e}]$ when $\gamma \Delta S > \beta \Delta d^2 \left[ F^*_T, \frac{n_e - 1}{n_e} \right] - \Delta c(F^*_T)$, and holds for $L_{TJ} \in [0, (L_{TJ})^{\dagger*}]$ when $\gamma \Delta S \in \left( \beta \Delta d^2 \left[ F^*_T, 0 \right] - \Delta c(F^*_T), \beta \Delta d^2 \left[ F^*_T, \frac{n_e - 1}{n_e} \right] - \Delta c(F^*_T) \right)$, where $(L_{TJ})^{\dagger*}$ is $L_{TJ}$ satisfying $\beta \Delta d^2 \left[ F^*_T, (n_e - 1) \left( \frac{L_{TJ}}{L} \right) \right] - \Delta c(F^*_T) = \gamma \Delta S$.

The second inequality of (82) holds for any $L_{TJ} \in [0, \frac{L}{n_e}]$ when $\gamma \Delta S < \beta \Delta d^2 \left[ F^*_T, 0 \right] - \Delta c(F^*_T)$, and for $L_{TJ} \in ((L_{TJ})^{\dagger*}, \frac{L}{n_e})$ when $\gamma \Delta S \in \left( \beta \Delta d^2 \left[ F^*_T, 0 \right] - \Delta c(F^*_T), \beta \Delta d^2 \left[ F^*_T, \frac{n_e - 1}{n_e} \right] - \Delta c(F^*_T) \right)$, where
(LTJ) (LTJ)^2 \text{ from } F_n^* < F_d(LTJ) \text{ is } LTJ \text{ satisfying } \beta \Delta d^2 \left[ F_d(LTJ), (n_e - 1) \frac{LTJ}{L} \left(2 - n_e \frac{LTJ}{L}\right) \right] - \Delta c(F_d(LTJ)) = \gamma \Delta S.

[Sector M] Because workers in sector M identify with the nation, the condition for them not to deviate from the equilibrium is given by (72) as in equilibrium (n). In the symmetric equilibrium, the condition becomes

\[
\beta \left[ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} - \omega_n n_e (n_e - 1) \left( \frac{LTJ}{L} \right)^2 \right] - \frac{1}{\theta} \left( \frac{n_e - 1}{n_e} \right)^{\theta - 1} \left[ (\delta V)^{\theta - 1} - (\delta V - \beta \eta)^{\theta - 1} \right] < \gamma \Delta S,
\]

where \(F\) is the solution for (32) and (33). When (82) and thus (32) hold, this condition holds.

[The equilibrium condition] Hence, when \(LTJ\) is the solution for (25), the equilibrium exists for \(LTJ \in ((LTJ)^2, (LTJ)^2)\) when \(\gamma \Delta S \in \left( \beta \Delta d^2 [F_n^*, 0] - \Delta c(F_n^*), \beta \Delta d^2 \left[ F_n^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_n^*) \right)\), and exists for \(LTJ \in (LTJ)^2, \frac{L}{n_e}\) when \(\gamma \Delta S \in \left( \beta \Delta d^2 [F_n^*, \frac{n_e - 1}{n_e}] - \Delta c(F_n^*), \beta \Delta d^2 \left[ F_e^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*) \right)\).