

MPRA

Munich Personal RePEc Archive

Modernization, Social Identity, and Ethnic Conflict

Yuki, Kazuhiro

September 2015

Online at <https://mpra.ub.uni-muenchen.de/67316/>
MPRA Paper No. 67316, posted 18 Oct 2015 18:14 UTC

Modernization, Social Identity, and Ethnic Conflict

Kazuhiro Yuki*

This version: September 2015

First version: March 2015

Abstract

Empirical evidence suggests that ethnic divisions or diversity in a society leads to negative outcomes in various dimensions, including civil conflict and economic development. 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. If shared national identity is important, how can it be realized? In political science, there exist conflicting theses emphasizing effects of modernization on national identity. Which thesis is more relevant under what conditions? How are conflict and output affected by modernization through identity? How do policies such as "nation-building" policies affect the outcome?

In order to examine these questions theoretically, this paper develops a model of social identity, ethnic conflict, and development. In the model, individuals choose a sector to work (between the modern sector and a traditional sector), social identity (between ethnic identity and national identity), and contributions to ethnic conflict. Thus, modernization (and output), identity, and conflict interact with each other.

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. Helpful comments from participants at the ERF workshop on Macroeconomics are gratefully acknowledged. Financial support from JSPS through Grants-in-Aid for Scientific Research 10197395 is acknowledged.

1 Introduction

Empirical evidence suggests that ethnic divisions or diversity in a society leads to negative outcomes in various dimensions, including internal armed conflict (Esteban, Mayoral, and Ray, 2012), public goods provision (Miguel and Gugerty, 2005), and economic development (Montalvo and Reynal-Quero, 2005) among others.¹ 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).

If shared national identity is important, how can it be realized? Miguel (2004) and Collier (2009), based on case study and statistical analysis, argue that "nation-building" policies, such as the promotion of a national language and school education emphasizing common history, culture, and values, are effective in strengthening national identity.² Meanwhile, in political science, there exist competing theses emphasizing roles of modernization (including industrialization, the rise of wage labor, the diffusion of education, and urbanization) on national identity (see Robinson, 2014, for an overview). The traditional thesis, based on the past experience of Europe, is that modernization leads to widespread national identity at the expense of ethnic and other subnational identities (Deutsch, 1953; Weber, 1979; Gellner, 1983). By contrast, based on post-independent experience of Africa, another influential thesis argues that modernization rather breeds ethnic identification due to intensified competition over resources (Melson and Wolpe, 1970; Bates, 1983).³ If the traditional view is correct, policies promoting modernization might be more important than "nation-building" policies for shared national identity and good outcomes in conflict and development, while if the competing view is true, nation-building policies would be critical. Both hypotheses would contain some truth, but which is more relevant under what conditions? How are conflict and output affected by modernization through identity? How do policies such as "nation-building" policies affect the outcomes?

In order 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

¹Esteban, Mayoral, and Ray (2012), employing a specification based on Esteban and Ray (2011) and cross-country data, find that three indices of ethnic divisions, polarization, fractionalization, and the Gini-Greenberg index, are significantly (positively for the first two indices, negatively for the last one) related to conflict. Miguel and Gugerty (2005) show that fractionalization is negatively related to school funding and school infrastructure quality in two districts of rural Kenya. Montalvo and Reynal-Quero (2005), using cross-country data, find the evidence suggesting that polarization has a negative effect on economic growth through its negative effect on investment and positive effects on government consumption and civil conflict and fractionalization has a direct negative effect on growth.

²Miguel (2004) finds that two neighboring rural districts of Tanzania and Kenya, which largely shared geography, history, and colonial institutional legacy, exhibit a sharp difference in the relationship between ethnic diversity and local provision of public goods (school funds and infrastructures), negative and significant for the Kenyan district and positive and 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 of post-independent governments contributed to differences in the strength of national identity and the above-mentioned relationship of the two countries.

³There is the third view focusing on Africa that also denies the traditional view based on much higher degrees of ethnic diversity of African nations than European nations (Kedourie, 1970; Davidson, 1992). This view also stresses important roles colonial legacy plays in today's social identity in Africa.

work (between the modern sector and a traditional sector), social identity (between ethnic and national identities), and contributions to conflict. Thus, modernization (and output), identity, and conflict interact with each other.

Model: The analysis is based on a contest model (Hirshleifer, 1989; Garfinkel, 1990; and Skaperdas, 1992) in which multiple ethnic groups contest for exogenous resources. A finite number of individuals belong to one of the ethnic groups that are symmetric in every aspect. So the model is concerned with an ethnically heterogeneous society without a dominant ethnic group.

There are multiple sectors producing the private good, ethnically-segregated traditional sectors and the integrated modern sector. The traditional sectors correspond to sectors using traditional or primitive technologies in the real economy, such as traditional agriculture, the urban informal sector, and household production, and the modern sector corresponds to sectors using modern technologies such as modern manufacturing and services; and the former sectors or activities tend to be more ethnically segregated than the latter: traditional agriculture is operated in rural communities and typical jobs in the urban informal sector are neighborhood jobs in ethnically segregated communities. The production technology of the modern sector exhibits constant returns to its only input, labor, and the wage is determined competitively, whereas, as in many existing works, that of traditional sectors exhibits decreasing returns to labor and the wage is determined so that the product is equally shared among workers.⁴ This setting can generate, in a simplest manner, the situation facing actual developing countries that there are inefficiently many workers in traditional sectors and their shift to the modern sector raises aggregate output.

The ethnic groups contest for exogenous resources that yield group-specific club goods, such as public services and infrastructures benefiting a specific group.⁵ The proportion of the resources a particular group acquires equals the proportion of contributions to conflict by members of the group. The individual cost of conflict increases with the amount of contribution or "efforts" an individual makes, and, as in standard contest models, the level of conflict is measured by the total "efforts" in the society. The resources represent both material resources (such as natural resources) and a part of the governmental budget used for producing the group-specific goods. The model considers the situation in which the resource allocation over the groups is determined not by rules but by the consequences of violent conflict or non-violent conflict (such as rent-seeking activities).

As in Sambanis and Shayo (2013), the utility of an individual depends not only on (i) his material payoff, which is the wage minus the cost of conflict plus the benefit from the group-specific club good, but also negatively on (ii) *perceived distance from a social group* he identifies with (his ethnic group or the nation) and positively on (iii) the *status of the social group*, which are important factors affecting social identification and intergroup behaviors, according to influential theories in

⁴The technology of traditional sectors intends to capture the fact that labor productivity tends to fall with the amount of labor input due to limited arable land (traditional agriculture), limited capital available to credit constrained producers (urban informal sector), or a decreasing degree of task specialization of each family member (household production). The wage determination rule reflects the fact that typical production units are family-run farms/firms or households. Except results on output, qualitative results do not depend on the latter assumption.

⁵Main results are unchanged when the contested resources yield the private good. But the assumption that the resources yield group-specific club goods justifies more naturally the setting that conflict arises among ethnic groups.

social psychology.⁶ He perceives how close to or far from a social group with the distance between his attributes and average attributes of the group. The attributes are whether one belongs to (a) the nation or not, (b) particular ethnic groups or not, and (c) particular traditional sectors or not. Following Sambanis and Shayo (2013), the weight on the ethnic attributes (national attribute) in the distance is assumed to be increasing (decreasing) in the level of conflict, which implies that, when conflict becomes more intense, people care about ethnicity more (nationality less) in measuring distances from social groups. The status of a social group is given by the exogenous difference between the group's "value" or "importance" and that of reference or comparison groups (other nations when the group is the nation). The national status, for example, represents people's evaluations of the nation's standing or reputation, particularly compared to neighboring nations, in "soft" dimensions including culture, history, sports, and widely shared values (such as human rights and democracy) as well as in "hard" dimensions such as military strength and territory.

The utility function implies that, given that an individual identifies with a particular social group, his utility increases as the perceived distance from the group decreases. Since the distance depends on differences in the sectoral attributes, others things equal, he has an incentive to choose the same sector as the "average person" of the group. Coordinating the choice among group members has positive externality on each other's utility. However, social identification of an individual is not fixed. He can "choose" a group (his ethnic group or the nation) that brings him higher utility either because of higher material payoff, the shorter perceived distance, or the higher status.⁷ His identity might change if exogenous variables affecting his utility or choices by others alter. For example, as the level of conflict rises, individuals place a greater (smaller) weight on the ethnic attributes (national attribute) in the perceived distance, which could change their identities.⁸

Individuals play a two-stage game to maximize their utility. First, they decide which sector to work, which determines labor incomes and sectoral and aggregate production. Then, they choose a social group to identify with and a contribution to conflict simultaneously, which determines the level of conflict, the allocation of the resources over the groups, and individual utilities.

Results: Equilibria can be classified into two types, those in which individuals of the same ethnic group share the same identity and those they do not.⁹ There exist two *homogenous identity equilibria*: the one in which all individuals identify with their ethnic group and the one in which all identify with the nation. And there exist three *heterogenous identity equilibria*: the one in which

⁶The 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 another influential theory, social identity theory (Tajfel and Turner, 1986).

⁷There is considerable evidence suggesting that perceived distance and status affect social identity. For example, Manning and Roy (2010) find, for Great Britain, that nonwhite individuals, whose perceived distance from 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 (that is, lower status) countries assimilate faster into British identity.

⁸Evidence suggests that people's identities are affected by the intensity of conflict. For example, Rohner, Thoenig, and Zilibotti (2013), using individual, county-level and district-level data from Uganda, find that the proportion of those identifying with their ethnic group over the nation is higher in counties of the higher intensity of armed conflicts, after controlling for individual, ethnic, and spatial characteristics and employing instrumental variable estimation.

⁹As mentioned above, ethnic groups are assumed to be symmetric in every aspect. Hence, the paper focuses on equilibria in which choices of all groups are symmetric. Most of asymmetric equilibria are very difficult to analyze.

those in the modern sector (traditional sectors) identify with the nation (their group); the one in which those in the modern sector are divided over identities and all in traditional sectors identify with their group; and the one in which those in traditional sectors are divided over identities (and all in the modern sector identify with the nation). Thus, modern sector workers are more (less) likely to identify with the nation (their ethnic group) than 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 with the nation above their ethnic group, after controlling for education, urban residence, gender, and various group-level and country-level variables.

When the equilibria are compared for given parameters and exogenous variables, it is found that the level of conflict is lower, the share of modern sector workers is higher, and, under conditions that would hold for typical developing countries, total output of the private good and aggregate material payoff are higher, when the proportion of individuals identifying with the nation is higher. That is, national identity is associated with not only the lower level of ethnic conflict, which is shown in Sambanis and Shayo (2013), but also the higher modern sector share and higher output.

While the previous comparison is made for given parameters and exogenous variables, which equilibrium(a) exist(s) changes with their values. A simple dynamics is introduced into the model by supposing that one of exogenous variables, the (total factor) productivity of the modern sector, increases over time. The productivity growth raises the modern sector wage, induces the higher proportion of workers to choose the sector, and raises the sector's share in production. How does such modernization of the economy affect social identity, conflict, and aggregate output?

If the national status is at extremes, the society stays in the same equilibrium: when the status is very high (very low), all individuals always identify with the nation (their ethnic group) and the level of conflict is consistently low (high). Otherwise, when the status is relatively high (low), the society tends to shift from a heterogenous identity equilibrium, in which traditional sector workers are more likely to identify with their ethnic group than modern sector workers, to the one in which all workers identify with the nation (their ethnic group) and the level of conflict is low (high). The sectoral shift of workers associated with modernization shakes social identities in both sectors: modern sector workers become *less* attached to the national identity and traditional sector workers become less attached to the ethnic identity. When the status is relatively high (low), the effect on traditional (modern) sector workers determines the equilibrium shift and all become identified with the nation (their group). Although the productivity increase always raises the modern sector's share in production and employment, *given* the productivity level, the society tends to be in an equilibrium characterized by large (small) modern sector shares and, under conditions that would hold for typical developing nations, high (low) aggregate output (and aggregate material payoff), when the status is high (low). That is, having sufficiently high national status is crucial in achieving universal national identity, a low level of conflict, high modern sector shares (a high degree of modernization), and high aggregate output in the long run.

However, history or "luck" too is important, as long as the status is not at extremes. Given

parameters and exogenous variables including the status, multiple equilibria tend to exist and thus identity, conflict, and output differ depending on which equilibrium is realized. Suppose that an equilibrium realized initially is maintained in subsequent periods (if it 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 be in an equilibrium with relatively strong national identity and relatively good conditions in other dimensions subsequently.

Similar results hold for contested resources too when "low (high) status" of the above result is replaced with "large (small) amount of resources". Specifically, *given the status*, when the amount of resources is large (small), the society tends to shift from a heterogeneous identity equilibrium to the one in which all individuals identify with their ethnic group (the nation). Note that the contested resources represent both material resources (such as natural resources) and a part of the governmental budget for group-specific goods whose allocation over the groups is determined by the consequences of violent or non-violent conflict. Hence, the result suggests that weak political and economic institutions as well as the abundance of material resources is a hindrance to the good outcomes. Further, an exogenous change making common nationality more salient (and ethnic differences less salient) in the perceived distance too has effects similar to an increase in the status.

The results are consistent with the traditional thesis on effects of modernization on social identity, if the national status is high, contested resources are not abundant, institutions are good in quality, or common nationality is valued (and ethnic differences are not valued), otherwise, they are consistent with the competing thesis, as far as the relatively long term effect is concerned. In the latter case, policies improving institutional quality, raising the national status, or making shared nationality more salient (and ethnic differences less salient) in people's minds are crucial for good outcomes.¹⁰ As mentioned above, Miguel (2004) and Collier (2009), based on case study and statistical analysis, argue that "nation-building" policies, including the promotion of a national language and school education emphasizing common history, culture, and values, are effective in strengthening national identity. The model shows how these policies can reinforce national identity through raising the national status or making shared nationality more salient. There are empirical works suggesting negative effects of natural resources on civil conflict and development and works suggesting important effects of political and economic institutions on civil conflict, rent-seeking activities, and development. The model reveals a novel mechanism interacting with social identity that resources and institutions affect ethnic conflict and development.

Related literature: This paper belongs to the literature examining issues on ethnic conflicts using contest models. Recent contributions include Esteban and Ray (2008, 2011), Besley and Persson (2010, 2011), Caselli and Coleman (2013), and Sambanis and Shayo (2013). Most closely related is Sambanis and Shayo (2013), who develop the first formal model of social identification and ethnic conflict using a utility function inspired by theories in social psychology. This paper

¹⁰As mentioned above, the national status represents people's evaluations of its international standing or reputation, particularly compared to neighboring nations, in "soft" dimensions such as culture, history, sports, and widely shared values and in "hard" dimensions such as military strength and territory. Clearly, policies can affect some of them.

extends their model by modeling multiple production sectors and associated sectoral choices by workers (and production decisions of firms), in order to examine interactions among modernization, identity, conflict, and output. Besley and Persson (2010) also examines the relation between conflict and development but in connection with capacities of the state to raise revenue and provide services supporting markets. Esteban and Ray (2008) develop a theory explaining the reason why ethnic conflict, rather than class conflict, is salient, especially in societies with distinct economic inequalities. Caselli and Coleman (2013) provide a theory of the salience of ethnic conflict from a different perspective. Besley and Persson (2011) examine conditions under which civil war and repression by a group holding power on the opposition group occur. Esteban and Ray (2011) construct a model that precisely connects the level of conflict with three measures of ethnic divisions.

The paper also belongs to the literature examining interactions between identity and economic behaviors, including Akerlof and Kranton (2000, 2010), Shayo (2009), Benabou and Tirole (2011), and Bisin et al. (2011). Akerlof and Kranton (2000) pioneer formally modeling and examining effects of identity on economic behaviors and Akerlof and Kranton (2010) illustrate how various behaviors can be explained by their framework. Shayo (2009) constructs the basic framework on which Sambanis and Shayo (2013) and this paper are based by generalizing the framework of Akerlof and Kranton (2000) and applies it to examine the political economy of income redistribution. Benbou and Tirole (2011) develop a general model of identity management, applicable to both personal and social identities, in which individuals who are uncertain about self-concept make investment in identities, and provide explanations for wide-ranging empirical puzzles. Bisin et al. (2011) develop a dynamic model of identity formation in which children of a minority group receive an identity trait (either “mainstream”, i.e. accept the values of the majority, or “oppositional”) from parents or role models and decide the intensity with which they identify with the trait.

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, 2014). In order 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 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.

Organization of the paper: Section 2 presents the model and 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, and Section 3.4 analyzes effects of resources on the interactions. Section 4 concludes. Appendix A presents existence conditions for equilibria, and Appendix B contains proofs.

2 Model

Consider a contest model (Hirshleifer, 1989; Garfinkel, 1990; and Skaperdas, 1992) in which $n_e (\geq 2)$ ethnic groups contest for exogenous resources. The society is populated by a finite number N of

individuals who belong to one of the ethnic groups that are symmetric in every aspect (thus the population size of each group is N/n_e). So the model is concerned with an ethnically heterogeneous society without a dominant ethnic group.

Production: There are $n_e + 1$ sectors producing the private good, n_e ethnically-segregated traditional sectors TJ ($J = 1, 2, \dots, n_e$) and the ethnically-integrated modern sector M . The traditional sectors correspond to sectors using traditional or primitive technologies in the real economy, such as traditional agriculture, the urban informal sector, and household production, and the modern sector corresponds to sectors using modern technologies such as modern manufacturing and services;¹¹ and the former sectors or activities tend to be more ethnically segregated than the latter: traditional agriculture is operated in rural communities and typical jobs in the urban informal sector are neighborhood jobs in ethnically segregated communities.

The production functions of sectors TJ ($J = 1, 2, \dots, n_e$) and M are

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

$$Y_M = A_M \sum_{J=1}^{n_e} L_{MJ}, \quad (2)$$

where L_{TJ} and A_T are respectively the number of workers in sector TJ and the sector's total factor productivity (TFP), 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.) Sector TJ exhibits decreasing returns to labor, which intends to capture the fact that labor productivity tends to fall 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 degree of task specialization of each family member (household production).¹²

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.¹³ Thus, labor incomes in the sectors are

$$y_{TJ} = A_T(L_{TJ})^{\alpha-1}, \quad (3)$$

$$y_M = A_M. \quad (4)$$

This setting can generate, in a simplest manner, the situation facing actual developing countries that there are inefficiently many workers in traditional sectors and their shift to the modern sector

¹¹The urban informal sector is a part of the urban economy composed of small-scale businesses supplying basic services (small shops and vendors selling commodities and meals are clear examples) and basic manufacturing goods. Even today, sectors or production activities using traditional technologies is important in most developing nations. Although urbanization lowered the share of agricultural employment significantly, it did not raise the share of sectors using modern technologies greatly in many countries. According to OECD (2009), informal employment, defined as the sum of urban informal-sector employment and formal-sector one without social protection (such as social security benefits) accounts for the majority of non-agricultural employment in developing nations.

¹²This is because the number of tasks performed by each family member increases as more production activities shift from the market to the household.

¹³This assumption reflects the fact that typical production units of traditional sectors are family-run farms/firms or households. Except results on total output of the private good, qualitative results below do not depend on this.

raises aggregate output.¹⁴

Conflict: The ethnic groups contest for exogenous resources that yield group-specific club goods of value V , such as public services and infrastructures benefiting a particular group.¹⁵ The amount of resources each group acquires depends on 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,

$$\frac{V_J}{V} = \frac{F_J}{F} \text{ if } F > 0, \text{ and } = \frac{1}{n_e} \text{ if } F = 0, \quad (5)$$

where V_J is the resources acquired by group J ($J = 1, 2, \dots, n_e$), $F_J = \sum_{i \in J} f_i$ 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, which is called the level of conflict.¹⁶ The contested resources represent both material resources (such as natural resources) and a part of the governmental budget used for producing the group-specific goods. The model considers the situation in which the resource allocation over the groups is determined not by rules but by the consequences of violent conflict or non-violent conflict (such as rent-seeking activities), in which force, mass demonstrations, bribery, or lobbying are employed to influence the outcome.

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:

$$c(f_i) = \frac{1}{\theta} (f_i)^\theta, \quad \theta \geq 2. \quad (6)$$

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 material payoff positively, but also negatively on *perceived distance* from a *social group* he identifies with (either his ethnic group or the nation) and positively on the *status* of the social group, which are important factors affecting social identification and intergroup behaviors, according to influential theories in social psychology.

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

$$\pi_i = y_K - \frac{1}{\theta} (f_i)^\theta + \delta \frac{F_J}{F} V, \quad (7)$$

where δ is the value of the group-specific club good in units of the private good.

¹⁴In 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 results.

¹⁵Main results are unchanged when the contested resources yield the private good. But the assumption that the resources yield group-specific club goods justifies more naturally the setting that conflict arises among ethnic groups. Note, however, that theories by Esteban and Ray (2008) and Caselli and Coleman (2013) provide explanations for the salience of ethnic conflict in the real society even when groups contest for private goods.

¹⁶As may be inferred from the setting, conflict always occurs in equilibrium. The paper does not examine the question of why costly and sometimes destructive conflict ever occurs at all, on which the literature provides informational asymmetries and commitment problems as major explanations (see Blattman and Miguel, 2010, for a review.) Rather, it is interested in how the level of conflict interacts with social identity and modernization, among others.

Social groups are groups from which an individual chooses one group he identifies with, which are, as in Sambanis and Shayo (2013), his ethnic group and the nation N . That is, the set of social groups for individuals of ethnic group J is $\{J, N\}$.

Individual i who is characterized by three types of attributes perceives how close to or far from a social group with the distance between his attributes and average attributes of the group. The attributes are whether one belongs to (a) the nation or not, (b) particular ethnic groups or not, and (c) particular traditional sectors or not:

$$q_i^n = 1 \text{ if } i \in N, \quad q_i^n = 0 \text{ otherwise,} \quad (8)$$

$$q_i^J = 1 \text{ if } i \in J, \quad q_i^J = 0 \text{ otherwise, for } J = 1, 2, \dots, n_e, \quad (9)$$

$$q_i^{TJ} = 1 \text{ if } i \in TJ, \quad q_i^{TJ} = 0 \text{ otherwise, for } J = 1, 2, \dots, n_e. \quad (10)$$

For example, when he belongs to ethnic group 2 and works in sector M , $q_i^n = 1, q_i^2 = 1, q_i^J = 0$ for $J \neq 2$, and $q_i^{TJ} = 0$ for any J . The national and ethnic attributes are fixed, while the sectoral attributes are determined endogenously by sectoral choices of workers, which are described later.

The *perceived distance* between individual i and social group G ($G = J, N$), on which his utility depends negatively, is represented by¹⁷

$$d_{iG}^2 = \omega_n (q_i^n - q_G^n)^2 + \omega_e \sum_{J=1}^{n_e} (q_i^J - q_G^J)^2 + \omega_s \sum_{J=1}^{n_e} (q_i^{TJ} - q_G^{TJ})^2, \quad (11)$$

where q_G^n, q_G^J , and q_G^{TJ} are average values of the three attributes of the group, and $\omega_n, \omega_e, \omega_s \in (0, 1)$ are weights on the respective attributes and their sum equals 1.

Following Sambanis and Shayo (2013), the weight on the ethnic attributes ω_e (the national attribute ω_n) is assumed to be increasing (decreasing) in the level of ethnic conflict F :¹⁸

$$\omega_e = \eta_0 + \eta_1 F, \quad \eta_0 \geq 0, \quad \eta_1 > 0, \quad \eta_0 + \eta_1 F_{\max} < 1 - \omega_s, \quad (12)$$

$$\omega_n = 1 - \omega_e - \omega_s = 1 - \omega_s - (\eta_0 + \eta_1 F), \quad (13)$$

where F_{\max} is the maximum possible level of F , whose value is determined later. The specification implies that, when ethnic conflict becomes more intense, people care about the ethnic attributes more (the national attribute less) in measuring distances from social groups.

The utility of an individual also depends positively on the *status of social group* G ($G = J, N$) he identifies with, which is given by the difference between the group's subjective "value" or "importance" and the reference groups' one:¹⁹

¹⁷The concept of perceived distance is developed in cognitive psychology in studying how a person categorizes information that comes in to her (stimuli) (Nosofsky, 1986). Turner et al. (1987) apply the concept to the categorization by a person of people, including herself, into social groups, in constructing an influential social psychological theory, self-categorization theory. The theory tries to explain psychological basis of social identification.

¹⁸The weight on the sectoral attributes, ω_s , is assumed to be independent of F , because it is not clear how ω_s is affected by the intensity of ethnic conflicts: if most people are in ethnically-segregated traditional sectors, ω_s would probably increase with F , while if most are in the integrated modern sector, ω_s would decrease with F .

¹⁹Intergroup status differences are major factors affecting intergroup behaviors such as conflict and discrimination, according to social identity theory (Tajfel and Turner, 1986), an influential social psychological theory closely related to self-categorization theory (footnote 17; see Turner and Reynolds, 2001, for example, for similarities and differences of the two theories), which tries to explain collective behaviors mainly based on social identity.

$$S_G = \sigma_G - \sigma_{-G}, \quad (14)$$

where exogenous σ_G and σ_{-G} summarize all factors affecting the group's and the reference or comparison groups' absolute "value" or "importance". When $G = J$, the reference group is the other ethnic groups, and when $G = N$, it is other nations. Since the ethnic groups are assumed to be symmetric, $S_J = \sigma_J - \sigma_{-J} = 0$, while $S_N = \sigma_N - \sigma_{-N}$ is generally non-zero.²⁰

The exogenous national status S_N represents people's evaluations of the nation's international standing or reputation, particularly compared to neighboring nations, in "soft" dimensions such as culture, history, sports, and widely shared values (for example, human rights and democracy) as well as in "hard" dimensions such as military strength and territory. In order to simplify the analysis greatly, unlike Sambanis and Shayo (2013), the status does not depend on the group's total material payoffs (the sum of π_i) and thus is exogenous. Results would not be affected by taking into account the economic status, as long as its importance in the utility is not very large.

From these settings, as in Sambanis and Shayo (2013), the utility of individual i who identifies with social group G is given by

$$u_{iG} = \pi_i - \beta d_{iG}^2 + \gamma S_G, \quad \beta, \gamma > 0. \quad (15)$$

The utility function implies that, given that an individual identifies with a particular social group, his utility increases as the perceived distance from the group decreases. Since the perceived distance depends on differences in the sectoral attributes, others things equal, he has an incentive to choose the same sector as the "average person" of the group. Coordinating the choice among group members has positive externality on each other's utility.

Social identification of an individual, that is, which group he identifies with, is not fixed. He can "choose" a group (his ethnic group or the nation) that brings him higher utility either because of higher material payoff, the shorter perceived distance, or the higher status.²¹ His social identity might change if exogenous variables affecting his utility directly or indirectly through choices by others alter. For example, as the level of conflict rises, individuals place a greater weight on the ethnic attributes and a smaller weight on the national attribute in the perceived distance, which could change their social identities.²² Exact timing of their decisions is as follows.

²⁰The assumption on reference groups is made for simplicity. If reference groups of the nation include ethnic groups and vice versa, $S_N = \sigma_N - [\rho\sigma_{-N} + (1 - \rho)\sigma_J]$ ($\rho \in [0, 1]$) and $S_J = \sigma_J - [\rho\sigma_{-J} + (1 - \rho)\sigma_N] = (1 - \rho)(\sigma_J - \sigma_N)$. Results in Section 3 remain the same if " S_N " is replaced with " $S_N - S_J$ " ($= (2 - \rho)\sigma_N - \rho\sigma_{-N} - 2(1 - \rho)\sigma_J$) in Sections 3.3 and 3.4 (and Appendix A).

²¹There is considerable evidence suggesting that perceived distance and status affect social identity. For example, Manning and Roy (2010) find, for Great Britain, that nonwhites, whose perceived distance from 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 (that is, lower status) countries assimilate faster into a British identity.

²²There is evidence suggesting that social identities of individuals are affected by the intensity of conflict. 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 and episodes suggesting the lack of strong ethnic identities before the war. Rohner, Thoening, and Zilibotti (2013), using individual, county-level and district-level data from Uganda, find that the proportion of individuals identifying with their ethnic group over the nation is higher in counties of the higher intensity of armed conflicts, after controlling for individual, ethnic, and spatial characteristics and employing instrumental variable estimation. Further, Eifert,

Timing: Individuals play a two-stage game to maximize their utility. First, they decide which sector to work (sector TJ or sector M for individuals of ethnic group J), which in turn determines labor incomes in traditional sectors (y_{TJ}) and sectoral and aggregate production (Y_{TJ}, Y_M , and $Y \equiv Y_{TJ} + Y_M$). Then, that is, after L_{TJ} and L_{MJ} are determined, they choose a social group to identify with and the contribution to conflict f_i simultaneously, which determines the level of conflict F , the allocation of the contested resources V over the groups, and individual utilities.²³ The solution concept applied is the subgame perfect Nash equilibrium, thus the two-stage game can be solved by backward induction.²⁴

3 Results

There exist many subgame perfect Nash equilibria for the game, including ones in which different ethnic groups make different choices, which are generally difficult to analyze. Hence, the paper focuses on equilibria in which choices of all ethnic groups are symmetric. 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 have different identities. 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 which equilibrium(a) exists do change with values of exogenous variables. Taking into account this, Section 3.3 analyzes the main focus of the paper, interactions among modernization (the shift from traditional sectors to the modern sector), identity, conflict, and output. And Section 3.4 examines how the abundance of contested resources affects the interactions.

In order 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.

$$\text{Assumption 1: } \delta \frac{V}{N} > (\beta \eta_1)^{\frac{\theta}{\theta-1}} \left(\frac{n_e - 1}{n_e} \right)^{\frac{1}{\theta-1}}. \quad (16)$$

3.1 Homogenous identity equilibria

There exist two homogenous identity equilibria, the equilibrium in which all individuals identify with their ethnic group and the one in which all individuals identify with the nation. The former

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.

²³The timing of events reflects the fact that the choice between the two sectors made earlier in life largely determines the sector to work for most of life (because, in the real economy, the sectors tend to require different levels of education and different types of skills and be located in different places), while social identity is more likely to change over time, usually gradually (see footnote 21 for the evidence on immigrants), but sometimes in a short period of time triggered by events such as armed conflict and electoral competition (see footnote 22 for the evidence).

²⁴Sambanis 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 familiar and it seems to be easier to apply. Shayo (2009) too employs the standard Nash equilibrium to solve a one-shot game of social identity.

is examined first.

3.1.1 All individuals identify with their ethnic group

Consider the second stage of the game in which sectoral allocation of workers (L_{TJ} and L_{MJ}) are given. When individual i of ethnic group J ($J = 1, 2, \dots, n_e$) in sector M identifies with his ethnic group, he chooses the contribution to conflict f_i to maximize the following utility:

$$A_M - \frac{1}{\theta}(f_i)^\theta + \delta \frac{F_J}{F} V - \beta \omega_s \left(\frac{L_{TJ}}{N/n_e} \right)^2. \quad (17)$$

From the first-order condition,

$$f_i = f_{i,e} \equiv \left(\delta \frac{F_{-J}}{F^2} V \right)^{\frac{1}{\theta-1}}, \text{ where } F_{-J} \equiv F - F_J. \quad (18)$$

When he is in sector TJ instead, he chooses f_i to maximize

$$A_T (L_{TJ})^{\alpha-1} - \frac{1}{\theta}(f_i)^\theta + \delta \frac{F_J}{F} V - \beta \omega_s \left(1 - \frac{L_{TJ}}{N/n_e} \right)^2. \quad (19)$$

The solution for f_i is given by (18) as in the previous case.

Since all individuals identify with their ethnic group and the ethnic groups are symmetric, by substituting $F_{-J} = \frac{n_e-1}{n_e}F$ and $f_i = F/N$ into (18), the equilibrium level of conflict F_e^* is obtained:

$$F_e^* = \left(\delta \frac{n_e-1}{n_e} \frac{V}{N} \right)^{\frac{1}{\theta}} N \text{ from } F_e^* = \left(\delta \frac{n_e-1}{n_e} \frac{V}{F_e^*} \right)^{\frac{1}{\theta-1}} N. \quad (20)$$

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

$$\text{Assumption 2: } A_T \left(\frac{N}{n_e} \right)^{\alpha-1} + \beta \omega_s < A_M. \quad (21)$$

Then, the sectoral allocation of workers is determined so that choosing either sector is indifferent. From (17) and (19), the indifference condition is

$$A_T (L_{TJ})^{\alpha-1} - \beta \omega_s \left(1 - 2n_e \frac{L_{TJ}}{N} \right) = A_M, \quad (22)$$

which gives the unique solution $(L_{TJ})_e^* \in (0, \frac{N}{n_e})$ that decreases with A_M and increases with A_T .²⁵

²⁵The first derivative with respect to L_{TJ} of the LHS of (22) is $-(1-\alpha)A_T(L_{TJ})^{\alpha-2} + \beta\omega_s \frac{2n_e}{N}$, which equals $-\infty$ at $L_{TJ} = 0$ and equals 0 at $L_{TJ} = \left[\frac{(1-\alpha)A_T}{\beta\omega_s \frac{2n_e}{N}} \right]^{\frac{1}{2-\alpha}}$, and the second derivative equals $(2-\alpha)(1-\alpha)A_T(L_{TJ})^{\alpha-3} > 0$. Thus, from (21) and the fact that the LHS of (22) at $L_{TJ} = 0$ equals $+\infty$, there exists unique $L_{TJ} \in (0, \frac{N}{n_e})$ satisfying (22). The relations of $(L_{TJ})_e^*$ with A_M and A_T are straightforward from the shape of the LHS of (22).

3.1.2 All individuals identify with the nation

Consider the second stage of the game in which sectoral allocation of workers are given. When individual i of ethnic group J in sector M identifies with the nation, he chooses f_i to maximize the following utility (note $\omega_e = \eta_0 + \eta_1 F$):

$$A_M - \frac{1}{\theta}(f_i)^\theta + \delta \frac{F_J}{F} V - \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\left(\frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\} + \gamma (\sigma_N - \sigma_{-N}). \quad (23)$$

From the first-order condition ($f_i > 0$ from the assumption (16)),

$$f_i = f_{i,n} \equiv \left(\delta \frac{F_{-J}}{F^2} V - \beta \eta_1 \frac{n_e - 1}{n_e} \right)^{\frac{1}{\theta - 1}}, \text{ where } F_{-J} \equiv F - F_J. \quad (24)$$

When he is in sector TJ instead, he chooses f_i to maximize

$$A_T (L_{TJ})^{\alpha - 1} - \frac{1}{\theta} (f_i)^\theta + \delta \frac{F_J}{F} V - \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\left(1 - \frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\} + \gamma (\sigma_N - \sigma_{-N}), \quad (25)$$

whose solution is given by (24) as in the previous case.

Since all individuals identify with the nation and the groups are symmetric, by plugging $F_{-J} = \frac{n_e - 1}{n_e} F$ and $f_i = F/N$ into (24), the equilibrium level of conflict F_n^* is obtained as a solution for

$$F_n^* = \left[\frac{n_e - 1}{n_e} \left(\delta \frac{V}{F_n^*} - \beta \eta_1 \right) \right]^{\frac{1}{\theta - 1}} N. \quad (26)$$

In the first stage, the indifference condition for sectoral choices equals, from (23) and (25),

$$A_T (L_{TJ})^{\alpha - 1} - \beta \omega_s \left(1 - 2 \frac{L_{TJ}}{N} \right) = A_M, \quad (27)$$

which gives the unique solution $(L_{TJ})_n^* \in (0, (L_{TJ})_e^*)$ that decreases with A_M and increases with A_T .²⁶

3.1.3 Analysis

The following proposition compares the two homogenous identity equilibria for *given* parameters and exogenous variables in terms of the level of conflict, the sectoral distribution of individuals, and output of the private good. As explained in Section 3.3 and detailed in Appendix A, there do exist combinations of parameters and exogenous variables such that both equilibria exist.

Proposition 1 *Given parameters and exogenous variables, the following holds for two homogenous identity equilibria.*

- (i) *The level of conflict is lower when all identify with the nation, i.e. $F_n^* < F_e^*$.*
- (ii) *L_{TJ} and thus the proportion of workers in traditional sectors are lower when all identify with the nation, i.e. $(L_{TJ})_n^* < (L_{TJ})_e^*$.*

²⁶From the comparison of the LHS of (27) with that of (22) and the discussion in footnote 25, it is clear that, when (21) is assumed, the unique solution $(L_{TJ})_n^* \in (0, (L_{TJ})_e^*)$ that decreases with A_M and increases with A_T exists.

(iii) Total output of the private good Y is higher under the national identity if α (the parameter of the traditional sector production function) is below a certain level or if A_M is not very high (or A_T is not very low). The output is higher under the ethnic identity if α is above a certain level (higher than the previous threshold) and A_M is high (or A_T is low) enough.

Individuals contribute less to conflict and thus the level of conflict F is lower when they identify with the nation, because they take into account the undesirable effect of the conflict on the perceived distance from the "average member" of the nation, in choosing f_i : higher F raises the weight on ethnicity, ω_e , and lowers the weight on nationality, ω_n , thereby highlighting differences among members and raising the distance. The same result is shown in Sambanis and Shayo (2013) and is consistent with empirical evidence such as Eifert, Miguel, and Posner (2010) and Rohner, Thoenig, and Zilibotti (2013) (see footnote 22).

What is new is the effect on the sectoral distribution of individuals and total output. Given parameters and exogenous variables, L_{TJ} and thus the proportion of workers in traditional sectors are lower when they identify with the nation. Under the national identity, the utility when they choose the traditional sector of their ethnic group is lower for given L_{TJ} and thus the smaller proportion of them choose the sector, because the perceived distance from the "average national" rises by choosing the ethnically segregated sector over the integrated modern sector, whereas, under the ethnic identity, the perceived distance from the "average member" of the ethnic group falls (if $L_{TJ} > \frac{N}{2n_e}$, i.e., the majority is in the traditional sector) or rises less (if $L_{TJ} < \frac{N}{2n_e}$).

In this model, the sectoral allocation of workers is generally inefficient, i.e. it does not maximize total output of the private good, because labor income equals the average labor productivity, not the marginal productivity, in traditional sectors (note, $\alpha < 1$, decreasing returns to labor in the sectors) and the perceived distance distorts sectoral choices. The former leads to too many workers in traditional sectors, while the latter leads to too few workers in the sectors under the national identity and to too many (few) workers in the sectors under the ethnic identity when $L_{TJ} > (<) \frac{N}{2n_e}$.

If α is below a certain level, the first effect dominates and L_{TJ} is higher than the efficient level. In this case, total output is higher under the national identity because L_{TJ} is smaller and thus closer to the efficient level. The condition would be more relevant to developing nations, since low α implies strong decreasing returns in traditional sectors.²⁷ The same result holds for *any* α , if A_M is not very high (or A_T is not very low) and thus L_{TJ} is not very small so that the second effect is positive (leads to too high L_{TJ}) or negative but small under the ethnic identity. As shown in the proof, this is the case if more than a quarter of workers choose traditional sectors under the national identity, but it could be true with a much smaller proportion in the sectors. The condition seems to hold in typical developing countries, considering the fact that a majority work

²⁷Remember 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). These factors are important in developing countries.

in the sectors (see footnote 11).²⁸

To summarize, national identity is associated with not only the lower level of conflict but also the higher share of modern sector workers and, under conditions that would hold at least for typical developing countries, higher levels of total output of the private good and of aggregate material payoff (the value of private and public good consumption net of the cost of conflict). Note that the result on output and material payoff holds despite the model does not assume the plausible negative effect of conflict on the modern sector productivity. The result would be strengthened if such effect is considered.²⁹

3.2 Heterogenous identity equilibria

Now, equilibria in which individuals of the same ethnic group have different identities are examined. There exist three heterogenous identity equilibria, the equilibrium in which sector M (sector TJ) workers identify with the nation (their ethnic group), the one in which those in sector M are divided over identities and all in sector TJ identify with their ethnic group, and the one in which those in sector TJ are divided over identities (and all in sector M identify with the nation).

3.2.1 Sector TJ workers identify with their ethnic group and sector M workers identify with the nation

In the second stage of the game in which sectoral allocation of workers are given, workers in sector TJ identifying with their ethnic group choose f_i to maximize (19) and the solution is given by (18), while those in sector M identifying with the nation choose f_i to maximize (23) and the solution is given by (24).

Because the ethnic groups are symmetric, by substituting $F_{-j} = \frac{n_e-1}{n_e}F$ into (18) and (24), and plugging them into $F = f_{i,e}n_eL_{TJ} + f_{i,n}(N - n_eL_{TJ})$, the level of conflict F given L_{TJ} is obtained:

$$F = \left(\frac{n_e-1}{n_e} \right)^{\frac{1}{\theta-1}} \left[\left(\delta \frac{V}{F} \right)^{\frac{1}{\theta-1}} n_e L_{TJ} + \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{1}{\theta-1}} (N - n_e L_{TJ}) \right], \quad (28)$$

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

In the first stage, the indifference condition for sectoral choices equals, from (18), (19), (23), (24), and (28),

²⁸By contrast, Y is lower under the national identity, if α is above a critical level (higher than the previous one) and A_M is high (or A_T is low) enough that $L_{TJ} < \frac{N}{2n_e}$, because the negative second effect dominates and thus L_{TJ} is lower than the efficient level in both equilibria. However, even in this case, unless Y is much lower, aggregate material payoff would be *higher* under the national identity because of the lower cost of conflict.

²⁹The 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.

$$\begin{aligned}
& A_T(L_{TJ})^{\alpha-1} - \frac{1}{\theta} \left(\delta \frac{n_e-1}{n_e} \frac{V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta-1}} - \beta \omega_s \left(1 - \frac{L_{TJ}}{N/n_e} \right)^2 \\
&= A_M - \frac{1}{\theta} \left(\delta \frac{n_e-1}{n_e} \frac{V}{F_d(L_{TJ})} - \beta \eta_1 \frac{n_e-1}{n_e} \right)^{\frac{\theta}{\theta-1}} - \beta \left\{ [\eta_0 + \eta_1 F_d(L_{TJ})]^{\frac{n_e-1}{n_e}} + \omega_s n_e \left(\frac{L_{TJ}}{N} \right)^2 \right\} + \gamma (\sigma_N - \sigma_{-N}) \\
&\Leftrightarrow A_T(L_{TJ})^{\alpha-1} + \beta \left\{ [\eta_0 + \eta_1 F_d(L_{TJ})]^{\frac{n_e-1}{n_e}} + \omega_s \left[n_e \left(\frac{L_{TJ}}{N} \right)^2 - \left(1 - n_e \frac{L_{TJ}}{N} \right)^2 \right] \right\} \\
&\quad - \frac{1}{\theta} \left(\frac{n_e-1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] - \gamma (\sigma_N - \sigma_{-N}) = A_M, \tag{29}
\end{aligned}$$

which gives the unique solution $(L_{TJ})_d^* \in (0, \frac{N}{n_e})$, as proved in Appendix B. The equilibrium level of conflict, F_d^* , is obtained from the substitution of $(L_{TJ})_d^*$ into (28).

3.2.2 Those in sector M are divided over identities and all in sectors TJ identify with their ethnic group

In the second stage of the game, workers in sector TJ identifying with their ethnic group choose f_i to maximize (19) and the solution is given by (18), while those in sector M are indifferent between identifying with the nation, in which case f_i is chosen to maximize (23) and the solution is given by (24), and identifying with their ethnic group, in which case f_i is chosen to maximize (17) and the solution is given by (18).

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

$$\begin{aligned}
& A_M - \frac{1}{\theta} (f_{i,e})^\theta + \delta \frac{F_I}{F} V - \beta \omega_s \left(\frac{L_{TJ}}{N/n_e} \right)^2 \\
&= A_M - \frac{1}{\theta} (f_{i,n})^\theta + \delta \frac{F_I}{F} V - \beta \left[(\eta_0 + \eta_1 F)^{\frac{n_e-1}{n_e}} + \omega_s n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] + \gamma (\sigma_N - \sigma_{-N}) \tag{30}
\end{aligned}$$

$$\Leftrightarrow \beta \left[(\eta_0 + \eta_1 F)^{\frac{n_e-1}{n_e}} - \omega_s n_e (n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \frac{1}{\theta} \left(\frac{n_e-1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] = \gamma (\sigma_N - \sigma_{-N}), \tag{31}$$

where F satisfies

$$\begin{aligned}
& F = f_{i,n} P_{M,n} (N - n_e L_{TJ}) + f_{i,e} [n_e L_{TJ} + (1 - P_{M,n})(N - n_e L_{TJ})] \\
&= \left(\frac{n_e-1}{n_e} \right)^{\frac{1}{\theta-1}} \left\{ \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{1}{\theta-1}} P_{M,n} (N - n_e L_{TJ}) + \left(\delta \frac{V}{F} \right)^{\frac{1}{\theta-1}} [n_e L_{TJ} + (1 - P_{M,n})(N - n_e L_{TJ})] \right\}, \tag{32}
\end{aligned}$$

where $P_{M,n}$ is the proportion of sector M workers identifying with the nation. Since the LHS of (31) decreases with L_{TJ} and increases with F , F satisfying (31) increases with L_{TJ} .

The indifference condition for sectoral choices in the first stage is given by (22) and thus the same as the equilibrium in which all identify with their ethnic group from (19) and (30). Thus, the equilibrium level of L_{TJ} , $(L_{TJ})_{Md}^*$, equals $(L_{TJ})_e^*$, and the equilibrium level of conflict F_{Md}^* is obtained by substituting $(L_{TJ})_e^*$ into (31) and solving it for F .

3.2.3 Those in sector TJ are divided over identities and all in sectors M identify with the nation

In the second stage, those in sector M identifying with the nation choose f_i to maximize (23) and the solution is (24), while those in sector TJ are indifferent between identifying with the nation, in which case f_i is chosen to maximize (25) and the solution is given by (24), and identifying with their ethnic group, in which case f_i is chosen to maximize (19) and the solution is given by (18).

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

$$\begin{aligned} & A_T(L_{TJ})^{\alpha-1} - \frac{1}{\theta}(f_{i,e})^\theta + \delta \frac{F_I}{F} V - \beta \omega_s \left(1 - \frac{L_{TJ}}{N/n_e}\right)^2 \\ &= A_T(L_{TJ})^{\alpha-1} - \frac{1}{\theta}(f_{i,n})^\theta + \delta \frac{F_I}{F} V - \beta \left\{ \omega_e \frac{n_e-1}{n_e} + \omega_s \left[\left(1 - \frac{L_{TJ}}{N}\right)^2 + (n_e-1) \left(\frac{L_{TK}}{N}\right)^2 \right] \right\} + \gamma(\sigma_N - \sigma_{-N}) \end{aligned} \quad (33)$$

$$\Leftrightarrow \beta \left[(\eta_0 + \eta_1 F) \frac{n_e-1}{n_e} + \omega_s (n_e-1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N}\right) \right] - \frac{1}{\theta} \left(\frac{n_e-1}{n_e}\right)^{\theta-1} \left[\left(\delta \frac{V}{F}\right)^{\theta-1} - \left(\delta \frac{V}{F} - \beta \eta_1\right)^{\theta-1} \right] = \gamma(\sigma_N - \sigma_{-N}), \quad (34)$$

where F satisfies

$$\begin{aligned} F &= f_{i,n} [P_{TJ,n} n_e L_{TJ} + (N - n_e L_{TJ})] + f_{i,e} (1 - P_{TJ,n}) n_e L_{TJ} \\ &= \left(\frac{n_e-1}{n_e}\right)^{\frac{1}{\theta-1}} \left[\left(\delta \frac{V}{F} - \beta \eta_1\right)^{\frac{1}{\theta-1}} [P_{TJ,n} n_e L_{TJ} + (N - n_e L_{TJ})] + \left(\delta \frac{V}{F}\right)^{\frac{1}{\theta-1}} (1 - P_{TJ,n}) n_e L_{TJ} \right], \end{aligned} \quad (35)$$

where $P_{TJ,n}$ is the proportion of sector TJ workers identifying with the nation. F satisfying (34) decreases with L_{TJ} because the LHS of (34) 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 (35).

The indifference condition for sectoral choices in the first stage is given by (27) and thus the same as the equilibrium in which all identify with the nation from (23) and (33). The equilibrium level of L_{TJ} , $(L_{TJ})_{Td}^*$, equals $(L_{TJ})_n^*$, and the equilibrium level of conflict F_{Td}^* is obtained by substituting $(L_{TJ})_n^*$ into (34) and solving it for F .

3.2.4 Analysis

In all the heterogenous 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 traditional sectors identify with the nation, all in the modern sector do (and when some in the modern sector identify with their ethnic group, all in traditional sectors do).³⁰ This is because the modern sector is ethnically integrated (and traditional sectors are ethnically segregated) and thus, if only sectoral attributes are considered, the perceived distance of modern sector workers from the "average national" is smaller than the distance from the "average member" of their group and the opposite holds for traditional sector workers.³¹ The result is consistent with Robinson (2014), who, using individual-level survey data of sixteen African countries, finds that being employed in

³⁰That there do not exist equilibria in which individuals in the modern sector are less likely to identify with the nation than those in traditional sectors is formally shown in the proof of Proposition A2 of Appendix A.

³¹The total perceived distance of modern sector workers from the "average national" could be *greater* than from the "average member" of their ethnic group, because ethnic attributes increase the distance from the "average national". By contrast, the total perceived distance of traditional sector workers from the "average national" is *always* greater than from the "average member" of their ethnic group.

the modern sector is significantly and robustly associated with identifying with the nation above their ethnic group, after controlling for education, urban residence, gender, and group-level and country-level variables.^{32,33}

The following proposition compares the heterogenous identity equilibria, together with the homogenous identity equilibria, in terms of the level of conflict, the sectoral distribution of workers, and output of the private good, for given parameters and exogenous variables. There do exist combinations of parameters and exogenous variables such that multiple or even all equilibria exist, as explained in Section 3.3 and detailed in Appendix A.

Proposition 2 *Given parameters and exogenous variables, the following holds when three heterogenous identity equilibria, together with two homogenous identity equilibria, are compared.*

- (i) *The level of conflict is lower when the proportion of individuals identifying with the nation is higher, i.e. $F_n^* < F_{Td}^* < F_d^* < F_{Md}^* < F_e^*$.*
- (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 α is below a certain level or if A_M is not very high (or A_T is not very low). The relation is opposite if α is above a certain level (larger than the previous threshold) and A_M is high (or A_T is low) enough.*

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

A rough explanation for the result is that individuals identifying with the national contribute less to conflict because, as explained after Proposition 1, in choosing f_i , they take into account that higher F raises the perceived distance from the "average national" by highlighting differences among ethnic groups.

The preceding explanation of the result on the conflict level presumes that, among the heterogenous identity equilibria, the proportion of individuals identifying with the nation is highest when those in sector TJ are divided over identities (and all in sector M identify with the nation) and lowest when those in sector M are divided over identities (and all in sector TJ identify with their ethnic group). The result on the fraction of workers in traditional sectors, $(L_{TJ})_{Td}^* = (L_{TJ})_n^* <$

³²Robinson (2014) classify 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.

³³By contrast, Eifert, Miguel, and Posner (2010), based on 22 public opinion surveys in 10 African countries, find that being a farmer or fisherman, whom they classify as traditional sector workers, is negatively correlated with the ethnic identity. However, there is no option for the national identity in the surveys (other options are religious and class/occupational identities) and, unlike this paper and Robinson (2014), they classify those in the urban informal sector as formal sector workers.

$(L_{TJ})_d^* < (L_{TJ})_{Md}^* = (L_{TJ})_e^*$, confirms that this is the case. $(L_{TJ})_n^* < (L_{TJ})_d^* < (L_{TJ})_e^*$ holds because, in the heterogenous identity equilibrium, the proportion of those identifying with the nation, who gain less from choosing the traditional sector of their ethnic group, is higher (lower) than in the equilibrium where all share the ethnic (national) identity. $(L_{TJ})_{Td}^* = (L_{TJ})_n^*$ and $(L_{TJ})_{Md}^* = (L_{TJ})_e^*$ hold, because individuals who share the same identity (the national identity for the former case and the ethnic identity for the latter) are in both sectors and thus they must be indifferent between the sectors as in homogenous identity equilibria.

Finally, the result on total output of the final good is similar to Proposition 1 and can be explained as before. The discussion on the result after Proposition 1 suggests that, for typical developing economies, total output is generally higher as the proportion of individuals having the national identity is higher, although $Y_{Td}^* = Y_n^*$ and $Y_{Md}^* = Y_e^*$ are true. Further, the aggregate material payoff is *strictly* higher when the proportion of those identifying with the nation is higher because of the lower cost of conflict.³⁴

To summarize, the results for homogenous identity equilibria in Proposition 1 continue to hold when heterogenous identity equilibria too are considered: national identity is associated with not only the lower level of conflict but also the higher share of modern sector workers and, under conditions that would hold at least for typical developing countries, higher levels of total output of the private good and of aggregate material payoff (the value of private and public good consumption net of the cost of conflict). The results are consistent with the often-made argument (for example, Collier, 2009) that the lack of shared social identity, that is, 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.

3.3 Interactions among modernization, identity, conflict and output

The previous sections compared different equilibria for given parameters and exogenous variables, but which equilibrium(a) exists changes with values of exogenous variables, as examined in detail in Appendix A. Taking into account this result, this section analyzes the main focus of the paper, interactions among modernization, identity, conflict and output.

A simple dynamics is introduced into the model by supposing that the TFP of sector M , A_M , increases over time. The productivity growth raises the modern sector income, induces the higher proportion of workers to choose the sector, i.e. lowers L_{TJ} , and raises the sector's share in production. How does modernization driven by the productivity growth affect social identity, conflict, and aggregate output?³⁵ The next proposition, based on the propositions in Appendix

³⁴The total cost of conflict is (N_n and N_e are respectively numbers of those identifying with the nation and their ethnic group) $\frac{1}{\theta} [(f_{i,n})^\theta N_n + (f_{i,e})^\theta N_e] = \frac{1}{\theta} [(f_{i,n})^{\theta-1} f_{i,n} N_n + (f_{i,e})^{\theta-1} f_{i,e} N_e] = \frac{1}{\theta} \frac{n_e-1}{n_e} [(\delta \frac{V}{F} - \beta \eta_1) f_{i,n} N_n + \delta \frac{V}{F} f_{i,e} N_e] = \frac{1}{\theta} \frac{n_e-1}{n_e} [\delta V - \beta \eta_1 f_{i,n} N_n]$, where the second equality is from the first order conditions of utility maximization and $F_{-J} = \frac{n_e-1}{n_e} F$. The total cost decreases with N_n since $f_{i,n}$ increases with N_n from (24) and Proposition 2.

³⁵Note 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 experienced rapid urbanization without significant modernization.

A, shows that the effect differs depending on the status of the nation S_N .³⁶ Note that changes in other exogenous variables including a decrease in contested resources V have similar effects to an increase in S_N , as shown in Propositions 4 and 5.

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

- (i) *If the status of the nation S_N is very high (very low), all individuals always identify with the nation (their ethnic group) and the level of conflict F is consistently low (high).*
- (ii) *Otherwise, when S_N is relatively high (low), the society tends to shift from a heterogenous identity equilibrium to the one in which all individuals identify with the nation (their ethnic group). Given parameters and exogenous variables including S_N , multiple equilibria tend to exist and thus social identity, conflict, and output differ depending on which equilibrium is realized.*
- (iii) *For given A_M , when S_N is relatively high (low), the society tends to be in an equilibrium with low (high) F and L_{TJ} and, under conditions that would hold for typical developing countries, a high (low) level of aggregate output of the final good Y (and aggregate material payoff).*

If the status of the nation S_N is at extremes, the society stays in the same equilibrium: when the status is very high (very low), all individuals always identify with the nation (their ethnic group) and the level of conflict F is consistently low (high).

Otherwise, when S_N is relatively high (low), the society tends to shift from a heterogenous identity equilibrium, in which traditional sector workers are more likely to identify with their ethnic group than modern sector workers, to the equilibrium in which all individuals identify with the nation (their ethnic group) and the level of conflict is low (high). The sectoral shift of workers associated with modernization shakes social identities in both sectors: modern sector workers become *less* attached to the national identity and traditional sector workers become less attached to the ethnic identity. When S_N is relatively high (low), it is usually the case that the effect on traditional (modern) sector workers determines the equilibrium shift and all become identified with the nation (their group). Although the increase of A_M always lowers L_{TJ} and raises the modern sector's share in production, for *given* A_M , the society tends to be in an equilibrium characterized by large (small) modern sector shares in employment and production and, under conditions that would hold for typical developing nations, high (low) aggregate output of the final good Y (and aggregate material payoff), when S_N is high (low). That is, having sufficiently high national status is crucial in achieving universal national identity, a low level of conflict, high modern sector shares (a high degree of modernization), and high aggregate output in the long run.

However, history or "luck" too is important, as long as the status is not at extremes. Given parameters and exogenous variables including S_N , multiple equilibria tend to exist and thus social identity, conflict, and output differ depending on which equilibrium is realized.³⁷ (This is the reason

³⁶As mentioned in footnote 20, when reference groups of the nation (an ethnic group) include ethnic groups (the nation) as well as other nations (ethnic groups) in the measure of the national status S_N (the status of ethnic group J , S_J), results in Sections 3.3 and 3.4 (and Appendix A) remain the same if " S_N " is replaced with " $S_N - S_J$ ".

³⁷In the model of Sambanis and Shyao (2013) too, multiple equilibria exist when the status is not at extremes.

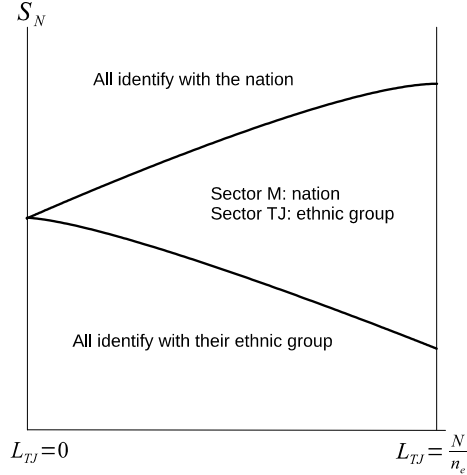


Figure 1: Equilibrium when $\eta_1 = 0$

for "tends to" in the statements of (ii) and (iii) of the proposition.) Suppose that an equilibrium realized initially is maintained in subsequent periods, if 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 be in an equilibrium with relatively strong national identity and relatively good conditions in terms of the level of conflict, modern sector shares, and aggregate output subsequently. This is particularly so when the status is at intermediate level, in which case multiple equilibria exist even in the long run (i.e., when A_M is high).

The result would be understood more easily, first by looking at the result when $\eta_1 = 0$, that is, when weights on ethnic attributes ω_e and on the national attribute ω_n of the perceived distance do not depend on the level of conflict F . In this case, F is the same in all equilibria and equilibrium is unique given parameters and exogenous variables. Figure 1 illustrates how the realized equilibrium differs depending on S_N and L_{TJ} , when $\eta_1 = 0$. As A_M increases over time, L_{TJ} decreases, whose value is determined by the indifference condition for the sectoral choice of the corresponding equilibrium. Thus, the society moves leftward with the productivity growth in the figure.

When the status of the nation is very high (very low), all individuals always identify with the nation (their ethnic group). By contrast, when the status is neither very high nor very low, the realized equilibrium changes with the productivity growth. When the status 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 identify with the nation (their ethnic group). That is, the social identity initially associated with modern (traditional) sector workers becomes the shared identity eventually, when the status is high (low). The growth of A_M raises the modern sector income and induces the higher proportion of workers to choose the sector. As a result, modern sector workers become *less* attached to the national identity (the difference in their utilities under the national identity and under the ethnic identity

falls), because the greater proportion of their ethnic group are in the sector and thus the perceived distance under the ethnic identity falls more than under the national identity,³⁸ while traditional sector workers become less attached to the ethnic identity, because the smaller proportion of their group are in the sector. That is, the sectoral shift of labor associated with modernization shakes long-standing identities in both sectors. When the national status is high, the latter effect on traditional sector workers determines the equilibrium shift (because utilities under the national identity are relatively high and thus the "identity shock" of modern sector workers are less severe) and all become identified with the nation, while when the status is low, the former effect on modern sector workers determines the shift and all become identified with their ethnic group.

When $\eta_1 > 0$, that is, when the weight on ω_e increases and the one on ω_n decreases with F , F is lower in an equilibrium with the higher proportion of individuals identifying with the nation (Propositions 1 and 2). Further, multiple equilibria could exist given parameters and exogenous variables and the two equilibria in which workers in one of the sectors are divided over identities could exist. Multiple equilibria could arise because of two-way positive causations between conflict and ethnic identity: when the level of conflict is high (low), people care about ethnicity more (less) in measuring the distance from social groups and thus they are more (less) likely to identify with their ethnic group, whereas when the proportion of those identifying with their ethnic group is high (low), the level of conflict is high (low). The two heterogeneous identity equilibria could exist because conflict depends on social identity when $\eta_1 > 0$: workers in one of the sectors can be indifferent between the identities in equilibrium, only if their identity choice affects F and thus utilities under the identities (note that (31) and (34) do not hold when $\eta_1 = 0$).

Figure 2 illustrates how the realized equilibrium(a) differs depending on S_N and L_{TJ} when $\eta_1 > 0$, based on Propositions A1 and A2 in Appendix A.³⁹ The equilibrium in which all identify with the nation (their ethnic group) exists in the region above the upward-sloping solid curve (on or below the downward-sloping solid curve). The equilibrium in which sector TJ workers identify with their ethnic group and sector M workers identify with the nation exists in the region with triple-dashed double-dotted lines (the lower borderline is not included), the one in which those in sector M are divided over identities and all in sectors TJ identify with their ethnic group exists in the region with negatively-sloped lines (both borderlines are not included), and the one in which those in sectors TJ are divided over identities and all in sector M identify with the nation exists in the region with positively-sloped lines (both borderlines are not included).

Suppose, for example, that the society starts with the equilibrium in which sector TJ workers identify with their ethnic group and sector M workers identify with the nation. As long as it stays in this equilibrium, with the growth of A_M , L_{TJ} and thus the proportion of individuals identifying with their ethnic group decrease, which results in a fall in the level of conflict, the increased share of

³⁸The perceived distance of modern sector workers falls under either identity but the fall under the ethnic identity is greater, because changes in the average sectoral attributes are greater.

³⁹The figure is for the case when ω_s is relatively high and η_1 is relatively low. Appendix A presents a figure when ω_s is relatively low and η_1 is relatively high (Figure 5). Although relative positions of several curves are different, basic features of the figure are similar to this one.

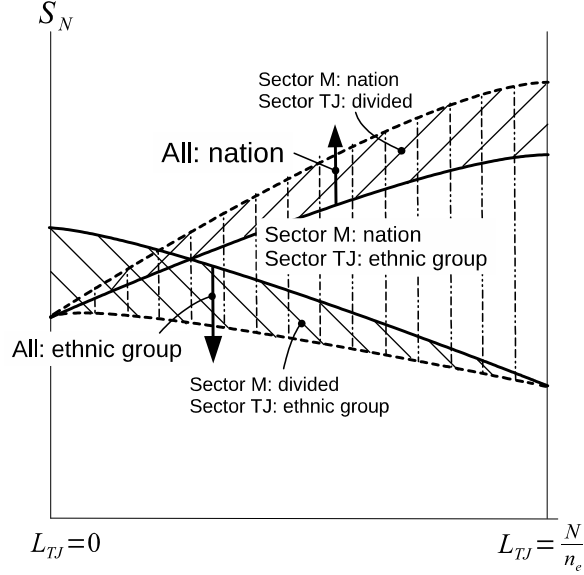


Figure 2: Equilibrium when $\eta_1 > 0$

the modern sector, and higher aggregate output and material payoff.^{40,41} However, this equilibrium ceases to exist and thus the society shifts to another equilibrium eventually. As the figure shows, if S_N is relatively high (low), it shifts to the equilibrium in which all individuals identify with the nation (their ethnic group) and the level of conflict falls (rises), where the rise of F when S_N is low may be interpreted as a rise in non-violent conflict such as rent-seeking activities, if it occurs at low L_{TJ} .⁴² The sectoral shift and output growth continue, but *given* A_M , L_{TJ} is lower, and the share of the modern sector, 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. The two equilibria in which workers in one of the sectors are divided over identities exist only in such regions. Suppose, for example, that the society starts with the region in which the following three equilibria – the equilibrium in which sector TJ workers identify with their ethnic group and sector M workers identify with the nation, the one in which all identify with their ethnic group, and the one in which sector M workers are divided over identities (and all in sectors TJ identify with their group) – exist. Depending on which equilibrium *happens to* be realized initially, social identity, the level of conflict, sectoral composition, and aggregate output in subsequent periods differ; clearly the outcome is worst when the society starts with the equilibrium in which all identify with their

⁴⁰ As mentioned above, it is supposed that an equilibrium realized initially is maintained in subsequent periods if the equilibrium continues to exist.

⁴¹ Actually, it is also possible that the proportion of individuals identifying with their ethnic group and the conflict level *increase* over time while L_{TJ} is relatively high. This is the case when the society is in the equilibrium in which those in sector TJ are divided over identities and all in sectors M identify with the nation. As mentioned just after (35), the proportion of those in sector TJ identifying with their ethnic group *increases* with a decrease in L_{TJ} .

⁴² The equilibrium shift changes F discontinuously, while it changes L_{TJ} continuously.

ethnic group. Further, when S_N is such that multiple equilibria exist even when L_{TJ} is low, the long-run outcome of the society starting with the equilibrium in which identity differs by sectors varies greatly depending on to which equilibrium the society *happens to shift*: the best (worst) is the shift to the one in which all identify with the nation (their ethnic group).

3.3.1 Discussion

As mentioned in Introduction, there exist competing theses on effects of modernization on social identity in political science. The traditional 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; Weber, 1979; Gellner, 1983), while another influential thesis mainly focusing on Africa argues that modernization rather breeds ethnic identification (Melson and Wolpe, 1970; Bates, 1983).

The proposition shows that, when the status of the nation is relatively high (low), eventually, the society tends to shift from a heterogenous identity equilibrium to the one in which all individuals identify with the nation (their ethnic group) characterized by a low (high) level of conflict. Thus, as far as the *relatively long term effect* of modernization (the effect involving the equilibrium shift) is concerned, the result is consistent with the traditional view when the status is relatively high, while it is consistent with the competing view when it is relatively low.

Robinson (2014), using cross-sectional individual-level survey data of sixteen African nations, finds that, when Tanzania is excluded from the sample, GDP per capita is significantly and positively related to identifying with the nation above 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. The evidence may be considered as capturing the *relatively short term effect* in an economy with a low degree of modernization, considering the fact that it is based on cross-sectional data of mostly poor African nations. Indeed, the effect of the productivity growth under a *given* equilibrium is consistent with her interpretation, when S_N is not at extremes and L_{TJ} is sufficiently high that the society is in a heterogenous identity equilibrium, except the one where sector TJ workers are divided over identities, in which case the effect is negative (see Figure 2).⁴³

However, both social identity and the degree of modernization for given modern sector productivity are *endogenously* determined in the model, thus her evidence can also partially be explained by the story that national identity and modernization are positively related through positive effects of the national status on them. Further, the evidence can partly be explained by multiple equilibria too, because national identity and modernization are positively related among different equilibria (for given parameters and exogenous variables). In order to distinguish the different stories empirically, it is important to estimate regression models with enough control variables using

⁴³In the equilibrium in which sector TJ workers are divided over identities, F increases with a decrease in L_{TJ} from (34). Then, the number of those identifying with their ethnic group, $(1 - P_{TJ,n})n_e L_{TJ}$, increases with a decrease in L_{TJ} from (35).

longitudinal data, although such data is not available at this point. Analysis using longitudinal data is also called for examining the relatively long term effect of modernization empirically.

The proposition also shows that multiple equilibria exist and the outcome depends on history or "luck" when the national status is not at extremes. As Sambanis and Shayo (2013) stress, this is consistent with the empirical finding that countries similar in ethnic diversity, geography, economic conditions, and political institutions have diverse histories regarding levels of ethnic conflict.

What are policy implications of the result that having sufficiently high national status is crucial in achieving the good outcome? The national status represents people's evaluations of its international standing or reputation, particularly compared to neighboring nations, in "soft" dimensions such as culture, history, sports, and widely shared values (for example, human rights and democracy) as well as in "hard" dimensions such as military strength and territory. Clearly, policies can affect some of the dimensions. Miguel (2004) and Collier (2009), based on case study and statistical analysis, argue that "nation-building" policies, including the promotion of a national language and school education emphasizing common history, culture, and values, are effective in strengthening national identity. The result shows how these policies can reinforce national identity through raising the national status and suggests that they are critical for the good outcome in countries where the status is low because of lack of shared culture and history, such as many African nations.

Alternatively, these "nation-building" policies may be interpreted as policies making shared nationality more salient, i.e. raising ω_n , and ethnicity less salient, i.e. reducing ω_e , in the perceived distance. Indeed, the following proposition shows that a decrease in η_0 that raises ω_n and lowers ω_e has similar effects to an increase in S_N . (Note that $\omega_e = \eta_0 + \eta_1 F$ and $\omega_n = 1 - \omega_s - (\eta_0 + \eta_1 F)$.)

Proposition 4 *Suppose that A_M increases over time. Then, results similar to Proposition 3 hold for η_0 when "high (low) S_N " is replaced with "low (high) η_0 ". Specifically, when η_0 is low (high), the society tends to shift from a heterogenous identity equilibrium to the one in which all individuals identify with the nation (their ethnic group), given other things equal.*

Graphically, this result holds because all the dividing lines of Figure 2 shift downward when η_0 decreases.

Empirical importance of "nation-building" policies on social identity seems to be supported. Miguel (2004) bases his argument on findings from comparison of Tanzania and Kenya, which largely shared geography, history, and colonial institutional legacy, but conducted sharply different ethnic policies after independence in areas such as national language and public school education and exhibits large differences in the strength of national identity. Robinson (2014) finds that the relation between modernization and national identity becomes insignificant once Tanzania is included into the sample because the national identity of Tanzania is "too strong" for given levels of control variables. More direct evidence is obtained for a developed country: Clots-Figueras and Masella (2013) examine effects of the introduction of a bilingual (Catalan and Spanish) education system in Catalonia using survey data and find that the number of years exposed to the bilingual system is positively related to the strength of Catalan identity and the propensity to vote for a party with a Catalanist platform.

3.4 Effects of contested resources

Finally, effects of the amount of contested resources V are examined. Specifically, how does it affect the level of conflict F and how does it influence the effects of the productivity growth on identity, conflict, and development? The following proposition summarizes the results.

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 "low (high) S_N " is replaced with "large (small) V ". Specifically, when V is large (small), the society tends to shift from a heterogenous identity equilibrium to the one in which all individuals identify with their ethnic group (the nation).*

The first result, which is consistent with many empirical studies on armed internal conflict such as Collier and Hoeffler (2004) and Ross (2006), 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 for resources V , when "low (high) S_N " of the proposition is replaced with "large (small) V ". Specifically, *given* the national status, when the amount of resources is large (small), the society tends to shift from a heterogenous identity equilibrium to the equilibrium in which all individuals identify with their ethnic group (the nation) with the productivity growth of the modern sector. That is, the abundance of resources is an impediment for the society to achieve universal national identity, a low level of conflict, a high share of the modern sector, and high aggregate output. Remember that the contested resources represent both material resources (such as natural resources) and a part of the governmental budget for group-specific club goods whose allocation over the groups is determined by the consequences of violent conflict or non-violent conflict (such as rent-seeking activities). Hence, the result shows that weak political and economic institutions as well as the abundance of material resources is a hindrance to the desirable outcome. The result is consistent with the traditional thesis on effects of modernization on social identity if contested resources are not abundant or 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 2 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. ω_e increases, (nationality less, i.e. ω_n decreases) in measuring perceived distances from social groups. Given the national status, this makes identifying with their ethnic group relatively more attractive compared to 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 that resources and institutions affect ethnic conflict and development.

4 Conclusion

Empirical evidence suggests that ethnic divisions or diversity in a society leads to negative outcomes in various dimensions, including civil conflict and economic development. 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. If shared national identity is important, how can it be realized? In political science, there exist conflicting theses emphasizing effects of modernization on national identity, the traditional thesis claiming the positive effect and the competing one claiming the negative effect. Which thesis is more relevant under what conditions? How are conflict and output affected by modernization through identity? How do policies such as "nation-building" policies affect the outcome?

In order 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 (between the modern sector and a traditional sector), social identity (between ethnic identity and national identity), and contributions to ethnic conflict. Thus, modernization (and output), identity, and conflict interact with each other.

It has been found that, given other things equal, a society with higher national status, less contested resources, better institutions, or ethnic differences less salient in people's minds tends to be in an equilibrium with the higher proportion of individuals with the national identity, the lower level of conflict, the higher share of modern sector workers (higher degree of modernization), and higher output. Simple dynamic analysis has shown that, as modernization proceeds, a society tends to shift to an equilibrium with uniformly national identity and the good outcome in other dimensions, if the status is high, the resources are not abundant, institutions are good in quality, or ethnic differences are not salient in people's minds; otherwise, it tends to shift to an equilibrium with uniformly ethnic identity and the worse outcome in other dimensions. Hence, the model is consistent with the traditional (competing) thesis on effects of modernization on identity under the former (latter) situation, as far as the relatively long term effect is concerned. When the national status is low or ethnic differences are salient in people's minds, "nation-building" policies, such as the promotion of a national language and school education emphasizing common history, culture, and values, are crucial for the good outcome. The model has shown how they can reinforce national identity through raising the national status or making shared nationality more salient. The model has also revealed a novel mechanism interacting with social identity that resources and institutions affect ethnic conflict and development.

References

- [1] Akerlof, G. A. and R. E. Kranton (2000), "Economics and Identity," *Quarterly Journal of Economics* 115 (3): 715-53.
- [2] Akerlof, G. A. and R. E. Kranton (2010), *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*, Princeton, New Jersey: Princeton University Press.
- [3] Banerjee, A. and A. Newman (1998), "Information, the dual economy, and development," *Review of Economic Studies* 65, 631-53.
- [4] Bates, R. (1983), "Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa." In *State Versus Ethnic Claims: African Policy Dilemmas*, ed. D. Rothchild and V. A. Olorunsola. London, UK: Westview Press, 152-171.
- [5] Benabou, R. and J. Tirole (2011), "Identity, morals, and taboos: Beliefs as assets," *Quarterly Journal of Economics* 126 (2): 805-855.
- [6] Besley, T. and T. Persson (2010), "State Capacity, Conflict, and Development," *Econometrica* 78 (1): 1-34.
- [7] Besley, T. and T. Persson (2011), "The Logic of Political Violence," *Quarterly Journal of Economics* 126 (3), 1411-1446.
- [8] Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2011), "Formation and persistence of oppositional identities," *European Economic Review* 55(8), 1046-1071.
- [9] Blattman, C. and E. Miguel (2010), "Civil War", *Journal of Economic Literature* 48 (1), 3-57.
- [10] Caselli, F. and W. J. Coleman II (2013), "On the Theory of Ethnic Conflict," *Journal of the European Economic Association* 11(S1): 161-192.
- [11] Clots-Figueras, I. and P. Masella (2013), "Education, Language and Identity," *Economic Journal* 123, F332-F357.
- [12] Collier, P. (2009), *Wars, Guns, and Votes*. New York, NY: Harper Perennial.
- [13] Collier, P. and A. Hoeffler (2004), "Greed and Grievance in Civil War," *Oxford Economic Papers* 56, 563-595.
- [14] Davidson, B. (1992), *The Black Man's Burden: Africa and the Curse of the Nation-State*. Oxford, UK: James Currey Publishers.
- [15] Deutsch, K. W. (1953), *Nationalism and Social Communication: An Inquiry into the Foundations of Nationality*. Cambridge, MA: MIT Press.
- [16] Easterly, W. (2001) "Can Institutions Resolve Ethnic Conflict?" *Economic Development and Cultural Change* 49(4), 687-706.
- [17] Eifert, B., E. Miguel, and D. N. Posner (2010), "Political Competition and Ethnic Identification in Africa," *American Journal of Political Science* 54(2): 494-510.
- [18] Esteban, J. and D. Ray (2008), "On the Saliency of Ethnic Conflict." *American Economic Review* 98 (5): 2185-202.

- [19] Esteban, J. and D. Ray (2011), "Linking Conflict to Inequality and Polarization," *American Economic Review* 101 (4): 1345–1374.
- [20] Esteban, J., L. Mayoral, and D. Ray (2012), "Ethnicity and Conflict: An Empirical Study," *American Economic Review* 102(4): 1310–42.
- [21] Garfinkel, M. R. (1990), "Arming as a Strategic Investment in a Cooperative Equilibrium," *American Economic Review* 80(1): 50–68.
- [22] Gellner, E. (1983), *Nations and Nationalism*. Ithaca, NY: Cornell University Press.
- [23] Hirshleifer, J. (1989), "Conflict and Rent-Seeking Success Functions: Ratio vs. Difference Models of Relative Success," *Public Choice* 63(2): 101–12.
- [24] Kedourie, Elie, ed. (1970), *Nationalism in Asia and Africa*. New York, NY: New American Library.
- [25] Lewis, A. W. (1954), "Economic development with unlimited supplies of labour," *Manchester School of Economic and Social Studies* 22 (2), 139–191.
- [26] Manning, A., and S. Roy (2010), "Culture Clash or Culture Club? National Identity in Britain," *Economic Journal* 120: F72–F100.
- [27] Mehlum, H., K. Moene, and R. Torvik (2006) "Institutions and the Resource Curse," *Economic Journal*, 116(508): 1–20.
- [28] Melson, R. and H. Wolpe (1970), "Modernization and the Politics of Communalism: A Theoretical Perspective," *American Political Science Review* 64(4): 1112–1130.
- [29] Miguel, E. (2004), "Tribe or Nation? Nation Building and Public Goods in Kenya Versus Tanzania," *World Politics* 56: 327–62.
- [30] Miguel, E. and M. K. Gugerty (2005), "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics* 89 (11-12): 2325–68.
- [31] Montalvo, J. G. and M. Reynal-Quero (2005), "Ethnic diversity and economic development," *Journal of Development Economics* 76 (2): 293–323.
- [32] Nosofsky, R. M. (1986), "Attention, Similarity and the Identification-Categorization Relationship," *Journal of Experimental Psychology: General* 115 (1): 39–57.
- [33] OECD (2009), *Is Informal Normal? Towards More and Better Jobs*, OECD, Paris.
- [34] Proto, E. (2007), "Land and the transition from a dual to a modern economy," *Journal of Development Economics* 83, 88–108.
- [35] Reynal-Querol, M. (2002) "Ethnicity, Political Systems, and Civil Wars," *Journal of Conflict Resolution* 46 (1), 29–54.
- [36] Robinson, A. L. (2014), "National Versus Ethnic Identification in Africa: Modernization, Colonial Legacy, and the Origins of territorial Nationalism," *World Politics* 66, 709–746.
- [37] Rodrik, D., A. Subramanian, and F. Trebbi (2004) "Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development," *Journal of Economic Growth* 9 (2), 131–165.

- [38] Rohner, D., M. Thoenig, and F. Zilibotti (2013), "Seeds of distrust: conflict in Uganda", *Journal of Economic Growth* 18: 217–252.
- [39] Ross, M. L. (2006) "A Closer Look at Oil, Diamonds, and Civil War," *Annual Review of Political Science*, 9: 265–300.
- [40] Sambanis, N. and M. Shayo (2013), "Social Identification and Ethnic Conflict," *American Political Science Review* 107 (2): 294–325.
- [41] Shayo, M. (2009), "A Model of Social Identity with an Application to Political Economy: Nation, Class and Redistribution," *American Political Science Review* 103 (2): 147–74.
- [42] Skaperdas, S. (1992), "Cooperation, Conflict, and Power in the Absence of Property Rights," *American Economic Review* 82(4): 720–39.
- [43] Tajfel, H. and J. C. Turner (1986), "The Social Identity Theory of Intergroup Behavior." In S. Worchel and W. Austin eds., *Psychology of Intergroup Relations*, Chicago: Nelson Hall, 7–24.
- [44] Turner, J. C., M. Hogg, P. Oakes, S. Reicher, and M. Wetherell (1987), *Rediscovering the Social Group: A Self-Categorization Theory*. Oxford: Blackwell.
- [45] Turner, J. C. and K. H. Reynolds (2001), "The Social Identity Perspective in Intergroup Relations: Theories, Themes, and Controversies." In Brown, R. and S. L. Gaertner, eds., *Blackwell Handbook of Social Psychology: Intergroup processes* 3 (1): 133–152.
- [46] Vollrath, D. (2009), "The dual economy in long-run development," *Journal of Economic Growth* 14, 287–312.
- [47] Weber, E. (1979), *Peasants into Frenchmen : The Modernization of Rural France, 1870-1914*. Stanford, CA: Stanford University Press.
- [48] Yuki, K. (2007), "Urbanization, informal sector, and development," *Journal of Development Economics* 84, 76–103.
- [49] Yuki, K. (2008), "Sectoral shift, wealth distribution, and development," *Macroeconomic Dynamics* 12, 527–559.
- [50] Yuki, K. (2014), "Education, Inequality, and Development in a Dual Economy," *Macroeconomic Dynamics*, forthcoming.

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 Figures 1 and 2 in Section 3.

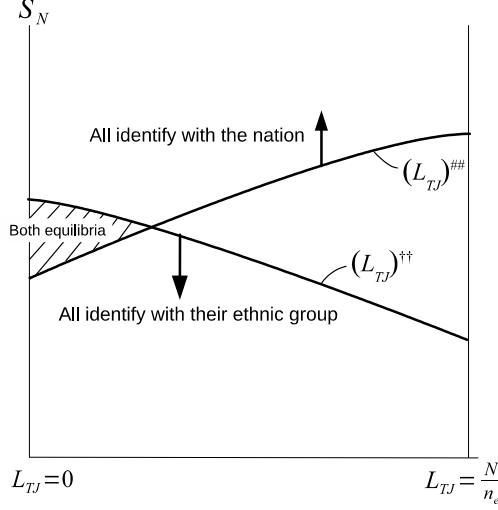


Figure 3: Proposition A1

A.1 Homogenous identity equilibria

The next proposition presents the existence conditions for the two homogenous identity equilibria. In the proposition, $\beta\Delta d^2[F, c_s] \equiv \beta \left[(\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + c_s \omega_s \right]$ (c_s is a coefficient on ω_s), $\Delta c(F) \equiv \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \left[\left(\delta \frac{V}{F} \right)^{\frac{\theta}{\theta - 1}} - \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{\theta}{\theta - 1}} \right]$, and $S_N \equiv \sigma_N - \sigma_{-N}$.

Proposition A1 (i) *The equilibrium in which all individuals identify with their ethnic group exists for any L_{TJ} when $\gamma S_N \leq \beta \Delta d^2 \left[F_e^*, -\frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*)$, and for $L_{TJ} \in [0, (L_{TJ})^{\dagger\dagger}]$ when $\gamma S_N \in \left(\beta \Delta d^2 \left[F_e^*, -\frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*), \beta \Delta d^2 [F_e^*, 0] - \Delta c(F_e^*) \right)$, where $L_{TJ} = (L_{TJ})_e^*$ is the solution for (22) and $(L_{TJ})^{\dagger\dagger}$ is the one for $\beta \Delta d^2 \left[F_e^*, -n_e(n_e - 1) \left(\frac{(L_{TJ})^{\dagger\dagger}}{N} \right)^2 \right] - \Delta c(F_e^*) = \gamma S_N$.*

(ii) *The equilibrium in which all identify with the nation exists for any L_{TJ} when $\gamma S_N > \beta \Delta d^2 \left[F_n^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_n^*)$, and for $L_{TJ} \in [0, (L_{TJ})^{##})$ when $\gamma S_N \in \left(\beta \Delta d^2 [F_n^*, 0] - \Delta c(F_n^*), \beta \Delta d^2 \left[F_n^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_n^*) \right)$, where $L_{TJ} = (L_{TJ})_n^*$ is the solution for (27) and $(L_{TJ})^{##}$ is the one for $\beta \Delta d^2 \left[F_n^*, (n_e - 1) \frac{(L_{TJ})^{##}}{N} \left(2 - n_e \frac{(L_{TJ})^{##}}{N} \right) \right] - \Delta c(F_n^*) = \gamma S_N$.*

Based on the proposition, Figure 3 illustrates combinations of L_{TJ} and S_N under which the homogenous identity equilibria exist. In the figure, the equilibrium in which all identify with the nation exists in the region above the upward-sloping curve and the one in which all identify with their ethnic group exists in the region on or below the downward-sloping curve. That is, given L_{TJ} , the former (latter) equilibrium tends to exist when S_N is high (low), and given S_N , the two equilibria tend to exist when L_{TJ} is low, where L_{TJ} is determined by the indifference condition for sectoral choices of the corresponding equilibrium. Note that both equilibria exist in the region with slant lines.

A.2 Heterogenous identity equilibria

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

Proposition A2 (i) *The equilibrium in which workers in sectors TJ identify with their ethnic group and those in sector M identify with the nation exists iff $\beta\Delta d^2\left[F_d(L_{TJ}), -(n_e-1)n_e\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_d(L_{TJ})) < \gamma S_N \leq \beta\Delta d^2\left[F_d(L_{TJ}), (n_e-1)\frac{L_{TJ}}{N}\left(2-n_e\frac{L_{TJ}}{N}\right)\right] - \Delta c(F_d(L_{TJ}))$, where $F_d(L_{TJ})$ is the solution for (28) and increases with L_{TJ} and $L_{TJ} = (L_{TJ})_d^*$ is the solution for (29).⁴⁴*

(ii) *The equilibrium in which those in sector M are divided over identities and all in sectors TJ identify with their ethnic group exists iff $\beta\Delta d^2\left[F_d(L_{TJ}), -(n_e-1)n_e\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_d(L_{TJ})) < \gamma S_N < \beta\Delta d^2\left[F_e^*, -n_e(n_e-1)\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_e^*)$, where $L_{TJ} = (L_{TJ})_e^*$ is the solution for (22).⁴⁵*

(iii) *The equilibrium in which those in sectors TJ are divided over identities and all in sector M identify with the nation exists for $L_{TJ} \in ((L_{TJ})^\sharp, (L_{TJ})^\sharp\sharp)$ when $\gamma S_N \in \left(\beta\Delta d^2[F_n^*, 0] - \Delta c(F_n^*), \beta\Delta d^2\left[F_n^*, \frac{n_e-1}{n_e}\right] - \Delta c(F_n^*)\right)$, and for $L_{TJ} \in ((L_{TJ})^\sharp, \frac{N}{n_e}]$ when $\gamma S_N \in \left(\beta\Delta d^2\left[F_n^*, \frac{n_e-1}{n_e}\right] - \Delta c(F_n^*), \beta\Delta d^2\left[F_e^*, \frac{n_e-1}{n_e}\right] - \Delta c(F_e^*)\right)$, where $L_{TJ} = (L_{TJ})_n^*$ is the solution for (27), $(L_{TJ})^\sharp$ is the one for $\beta\Delta d^2\left[F_d((L_{TJ})^\sharp), (n_e-1)\frac{(L_{TJ})^\sharp}{N}\left(2-n_e\frac{(L_{TJ})^\sharp}{N}\right)\right] - \Delta c(F_d((L_{TJ})^\sharp)) = \gamma S_N$, and $(L_{TJ})^\sharp\sharp$ is the one for $\beta\Delta d^2\left[F_n^*, (n_e-1)\frac{(L_{TJ})^\sharp\sharp}{N}\left(2-n_e\frac{(L_{TJ})^\sharp\sharp}{N}\right)\right] - \Delta c(F_n^*) = \gamma S_N$.*

Based on the proposition, Figure 4 illustrates combinations of L_{TJ} and S_N under which each equilibrium including the homogenous identity equilibria exists when ω_s (the weight on sectoral attributes in the perceived distance) is relatively high and η_1 (the strength of the effect of F on the weights on ethnic and national attributes) is relatively low.⁴⁶ The equilibrium in which workers in sectors TJ identify with their ethnic group and those in sector M identify with the nation exists in

⁴⁴To be more detailed, the equilibrium exists for $L_{TJ} \in [(L_{TJ})^\sharp, \frac{N}{n_e}]$ when $\gamma S_N \in \left(\max_{L_{TJ}}\left\{\beta\Delta d^2\left[F_d(L_{TJ}), -(n_e-1)n_e\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_d(L_{TJ}))\right\}, \beta\Delta d^2\left[F_e^*, \frac{n_e-1}{n_e}\right] - \Delta c(F_e^*)\right)$ and for $L_{TJ} \geq \max\{(L_{TJ})^\sharp, 0\}$ satisfying $\beta\Delta d^2\left[F_d(L_{TJ}), -(n_e-1)n_e\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_d(L_{TJ})) < \gamma S_N$ when $\gamma S_N \in \left(\min_{L_{TJ}}\left\{\beta\Delta d^2\left[F_d(L_{TJ}), -(n_e-1)n_e\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_d(L_{TJ}))\right\}, \max_{L_{TJ}}\left\{\beta\Delta d^2\left[F_d(L_{TJ}), -(n_e-1)n_e\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_d(L_{TJ}))\right\}\right)$.

⁴⁵To be more detailed, the equilibrium exists for $L_{TJ} \in [0, (L_{TJ})^\sharp\sharp)$ when $\gamma S_N \in \left(\max_{L_{TJ}}\left\{\beta\Delta d^2\left[F_d(L_{TJ}), -(n_e-1)n_e\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_d(L_{TJ}))\right\}, \beta\Delta d^2[F_e^*, 0] - \Delta c(F_e^*)\right)$ and for $L_{TJ} < \min\{(L_{TJ})^\sharp\sharp, \frac{N}{n_e}\}$ satisfying $\beta\Delta d^2\left[F_d(L_{TJ}), -(n_e-1)n_e\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_d(L_{TJ})) < \gamma S_N$ when $\gamma S_N \in \left(\min_{L_{TJ}}\left\{\beta\Delta d^2[F_n^*, 0] - \Delta c(F_n^*), \beta\Delta d^2\left[F_e^*, -\frac{n_e-1}{n_e}\right] - \Delta c(F_e^*)\right\}, \max_{L_{TJ}}\left\{\beta\Delta d^2\left[F_d(L_{TJ}), -(n_e-1)n_e\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_d(L_{TJ}))\right\}\right)$, where $(L_{TJ})^\sharp\sharp$ is L_{TJ} satisfying $\beta\Delta d^2\left[F_e^*, -n_e(n_e-1)\left(\frac{L_{TJ}}{N}\right)^2\right] - \Delta c(F_e^*) = \gamma S_N$.

⁴⁶To be more accurate, this is the case when $\beta\Delta d^2\left[F_n^*, \frac{n_e-1}{n_e}\right] - \Delta c(F_n^*) > \beta\Delta d^2[F_e^*, 0] - \Delta c(F_e^*) \Leftrightarrow \beta\Delta d^2[F_n^*, 0] - \Delta c(F_n^*) > \beta\Delta d^2\left[F_e^*, -\frac{n_e-1}{n_e}\right] - \Delta c(F_e^*) \Leftrightarrow \beta\frac{n_e-1}{n_e}\omega_s > \beta\frac{n_e-1}{n_e}\eta_1(F_e^* - F_n^*) + [\Delta c(F_n^*) - \Delta c(F_e^*)]$ holds, where $\beta\Delta d^2\left[F_n^*, \frac{n_e-1}{n_e}\right] - \Delta c(F_n^*)$ ($\beta\Delta d^2[F_e^*, 0] - \Delta c(F_e^*)$) is the value of γS_N at the intersection of the upward-sloping

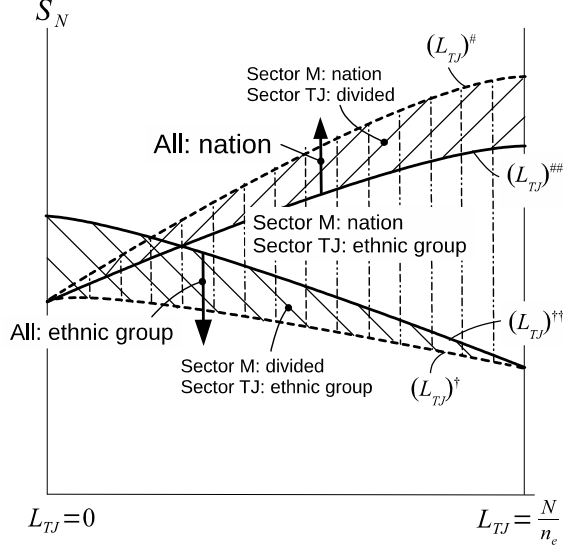


Figure 4: Proposition A2 when ω_s is relatively high and η_1 is relatively low

the region with triple-dashed double-dotted lines (the lower borderline is not included), the one in which those in sector M are divided over identities and all in sectors TJ identify with their ethnic group exists in the region with negatively-sloped lines (both borderlines are not included), and the one in which those in sectors TJ are divided over identities and all in sector M identify with the nation exists in the region with positively-sloped lines (both borderlines are not included). As in Figure 3, the equilibrium in which all identify with the nation (their ethnic group) exists in the region above the upward-sloping solid curve (on or below the downward-sloping solid curve).

Given L_{TJ} (which is determined by the indifference condition for sectoral choices of the corresponding equilibrium), the heterogenous identity equilibria tend to exist when S_N is neither very high nor very low, and given S_N , the equilibrium in which those in sectors TJ identify with their ethnic group and those in sector M identify with the nation tends to exist when L_{TJ} is large. Note that there are several regions in which multiple equilibria exist. In particular, the two equilibria in which workers in one of the sectors are divided over identities exist only in such regions. All five equilibria exist in the small triangular region on the left side of the figure.

Figure 5 illustrates combinations of L_{TJ} and S_N under which each equilibrium exists when ω_s is relatively low and η_1 is relatively high. Unlike Figure 4, the value of S_N at the intersection of the downward-sloping solid curve with $L_{TJ} = 0$ is greater than the one at the intersection of the upward-sloping solid curve with $L_{TJ} = \frac{n_e}{N}$, and the value of S_N at the intersection of the bottom dotted curve with $L_{TJ} = 0$ is smaller than the one at the intersection of the curve with $L_{TJ} = \frac{n_e}{N}$. However, basic features of the figure are similar to the previous one.

(downward-sloping) solid curve with $L_{TJ} = \frac{N}{n_e}$ ($L_{TJ} = 0$), and $\beta\Delta d^2[F_n^*, 0] - \Delta c(F_n^*)$ ($\beta\Delta d^2[F_e^*, -\frac{n_e-1}{n_e}] - \Delta c(F_e^*)$) is the value of γS_N at the intersection of the bottom dotted curve with $L_{TJ} = 0$ ($L_{TJ} = \frac{N}{n_e}$). The LHS of the last inequality increases with ω_s , while the RHS can be shown to increase with η_1 .

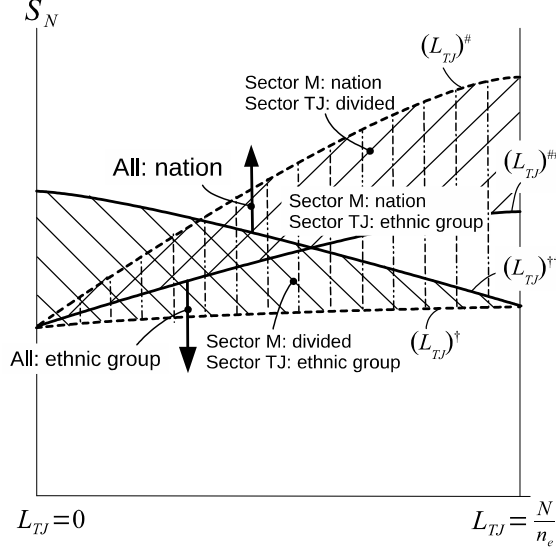


Figure 5: Proposition A2 when ω_s is relatively low and η_1 is relatively high

Appendix B Proofs

Proof of the uniqueness of $(L_{TJ})_d^*$. The derivative of the LHS of (29) with respect to L_{TJ} equals

$$\begin{aligned}
& -(1-\alpha)A_T(L_{TJ})^{\alpha-2} + 2\beta\omega_s \frac{n_e}{N} \left[1 - (n_e-1)\frac{L_{TJ}}{N} \right] \\
& + \frac{n_e-1}{n_e} \left\{ \beta\eta_1 + \frac{1}{\theta-1} \delta \frac{V}{(F_d(L_{TJ}))^2} \left(\frac{n_e-1}{n_e} \right)^{\frac{1}{\theta-1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{1}{\theta-1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta\eta_1 \right)^{\frac{1}{\theta-1}} \right] \right\} F'_d(L_{TJ}), \quad (36)
\end{aligned}$$

where, from (28),

$$F'_d(L_{TJ}) = \frac{\left(\frac{n_e-1}{n_e} \right)^{\frac{1}{\theta-1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{1}{\theta-1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta\eta_1 \right)^{\frac{1}{\theta-1}} \right] n_e}{1 + \frac{1}{\theta-1} \delta \frac{V}{(F_d(L_{TJ}))^2} \left(\frac{n_e-1}{n_e} \right)^{\frac{1}{\theta-1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{1}{\theta-1}-1} n_e L_{TJ} + \left(\delta \frac{V}{F_d(L_{TJ})} - \beta\eta_1 \right)^{\frac{1}{\theta-1}-1} (N - n_e L_{TJ}) \right]}. \quad (37)$$

The second derivative of the LHS of (29) with respect to L_{TJ} equals

$$\begin{aligned}
& (2-\alpha)(1-\alpha)A_T(L_{TJ})^{\alpha-3} - 2\beta\omega_s \frac{n_e(n_e-1)}{N^2} \\
& + \frac{d \left(\frac{n_e-1}{n_e} \left\{ \beta\eta_1 + \frac{1}{\theta-1} \delta \frac{V}{(F_d(L_{TJ}))^2} \left(\frac{n_e-1}{n_e} \right)^{\frac{1}{\theta-1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{1}{\theta-1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta\eta_1 \right)^{\frac{1}{\theta-1}} \right] \right\} F'_d(L_{TJ}) \right)}{dL_{TJ}}. \quad (38)
\end{aligned}$$

Since, as shown in the proof of Proposition 2 (iii), the LHS of (29) is always lower than that of (22), $(L_{TJ})_d^* < (L_{TJ})_e^*$ holds. Thus, the derivative of the LHS of (22) for $L_{TJ} \leq (L_{TJ})_d^*$ is negative, that is, $-(1-\alpha)A_T(L_{TJ})^{\alpha-2} + 2\beta\omega_s \frac{n_e}{N} < 0$. From this inequality, the first part of (38) is positive for $L_{TJ} \leq (L_{TJ})_d^*$, since it is greater than $(1-\alpha)A_T(L_{TJ})^{\alpha-3} \left[(2-\alpha) - L_{TJ} \frac{n_e-1}{N} \right] > 0$.

The second part of (38) is positive at $L_{TJ} \leq (L_{TJ})_d^*$ when $\theta = 2$, because the second part of (36)

equals

$$\frac{n_e-1}{n_e} \beta \eta_1 \left[1 + \delta \frac{V}{(F_d(L_{TJ}))^2} \frac{n_e-1}{n_e} \right] \frac{\frac{n_e-1}{n_e} \beta \eta_1 n_e}{1 + \delta \frac{V}{(F_d(L_{TJ}))^2} \frac{n_e-1}{n_e} N}, \quad (39)$$

which clearly increases with L_{TJ} . Hence, the second derivative of the LHS of (29) is positive for $L_{TJ} \leq (L_{TJ})_d^*$ and thus $(L_{TJ})_d^*$ is unique (remember that the LHS of (29) is always lower than that of (22) whose solution is unique $(L_{TJ})_e^*$). This also proves that the first derivative of the LHS is negative for $L_{TJ} \leq (L_{TJ})_d^*$. ■

Proof of Proposition 1. (i) Straightforward from (20) and (26). (ii) Shown in footnotes 25 and 26.

(iii) Denote total output of the private good in the equilibrium in which all individuals identify with the nation (their ethnic group) by Y_n^* (Y_e^*). From (1) and (2),

$$Y_n^* > Y_e^* \Leftrightarrow A_T((L_{TJ})_n^*)^\alpha - A_M(L_{TJ})_n^* > A_T((L_{TJ})_e^*)^\alpha - A_M(L_{TJ})_e^* \quad (40)$$

$$\Leftrightarrow \left(1 - 2\frac{(L_{TJ})_n^*}{N}\right)(L_{TJ})_n^* - \left(1 - 2n_e\frac{(L_{TJ})_e^*}{N}\right)(L_{TJ})_e^* > 0 \text{ (from (22) and (27))} \quad (41)$$

$$\Leftrightarrow ((L_{TJ})_e^* - (L_{TJ})_n^*) \left[\frac{2n_e}{N} ((L_{TJ})_e^* + (L_{TJ})_n^*) - 1 \right] + \frac{2}{N} (n_e - 1) ((L_{TJ})_n^*)^2 > 0. \quad (42)$$

Since $(L_{TJ})_e^* > (L_{TJ})_n^*$, the above condition holds when A_M is not very high or A_T is not very low so that $\frac{2n_e}{N} ((L_{TJ})_e^* + (L_{TJ})_n^*) - 1 > \frac{4n_e}{N} (L_{TJ})_n^* - 1 \geq 0$ is true, although the condition could hold even when $(L_{TJ})_n^*$ is much lower than $\frac{N}{4n_e}$ if the second term dominates the first term. For example, if parameters and exogenous variables are such that $(L_{TJ})_n^* \geq (L_{TJ})_e^* - (L_{TJ})_n^*$, then the inequality holds when $\frac{(L_{TJ})_n^*}{N} \geq \frac{1}{4n_e + 2(n_e - 1)}$.

The derivative of total output with respect to L_{TJ} equals $n_e [\alpha A_T (L_{TJ})^{\alpha-1} - A_M]$ from (1) and (2), which is negative (positive) when $L_{TJ} > (<) L_{TJ}^{**} \equiv \left(\frac{\alpha A_T}{A_M}\right)^{\frac{1}{1-\alpha}} (< \frac{N}{n_e}$ from the assumption (21)). Thus, $Y_n^* > Y_e^*$ holds if

$$(L_{TJ})_n^* > L_{TJ}^{**} \Leftrightarrow A_T (L_{TJ}^{**})^{\alpha-1} - \beta \omega_s \left(1 - 2\frac{L_{TJ}^{**}}{N}\right) = \frac{A_M}{\alpha} - \beta \omega_s \left[1 - 2\left(\frac{\alpha A_T}{A_M}\right)^{\frac{1}{1-\alpha}} \frac{1}{N}\right] > A_M \text{ from (27) and Assumption 2,} \quad (43)$$

where the derivative of the LHS of this equation with respect to α equals

$$-\frac{A_M}{\alpha^2} + \beta \omega_s \frac{2}{N} \frac{1}{1-\alpha} \left[\frac{1}{1-\alpha} \ln\left(\frac{\alpha A_T}{A_M}\right) + \frac{1}{\alpha} \right] \left(\frac{\alpha A_T}{A_M}\right)^{\frac{1}{1-\alpha}} \quad (44)$$

and the second derivative equals

$$\begin{aligned} & 2\frac{A_M}{\alpha^3} + \beta \omega_s \frac{2}{N} \left(\frac{1}{1-\alpha} \left[\frac{2}{(1-\alpha)^2} \ln\left(\frac{\alpha A_T}{A_M}\right) + \frac{1}{\alpha} \left(\frac{2}{1-\alpha} - \frac{1}{\alpha}\right) \right] + \left\{ \frac{1}{1-\alpha} \left[\frac{1}{1-\alpha} \ln\left(\frac{\alpha A_T}{A_M}\right) + \frac{1}{\alpha} \right] \right\}^2 \right) \left(\frac{\alpha A_T}{A_M}\right)^{\frac{1}{1-\alpha}} \\ & = 2\frac{A_M}{\alpha^3} + \beta \omega_s \frac{2}{N} \frac{1}{1-\alpha} \left\{ \left[\frac{2}{(1-\alpha)^2} \ln\left(\frac{\alpha A_T}{A_M}\right) + \frac{1}{\alpha} \left(\frac{2}{1-\alpha} + \frac{\alpha}{1-\alpha} \frac{1}{\alpha}\right) \right] + \frac{1}{1-\alpha} \left[\frac{1}{1-\alpha} \ln\left(\frac{\alpha A_T}{A_M}\right) + \frac{2}{\alpha} \right] \frac{1}{1-\alpha} \ln\left(\frac{\alpha A_T}{A_M}\right) \right\} \left(\frac{\alpha A_T}{A_M}\right)^{\frac{1}{1-\alpha}} > 0, \end{aligned} \quad (45)$$

from $L_{TJ}^{**} = \left(\frac{\alpha A_T}{A_M}\right)^{\frac{1}{1-\alpha}} > 1$.

Because (43) holds as $\alpha \rightarrow 0$, does not hold as $\alpha \rightarrow 1$, and the first derivative of the LHS of the equation is $-\infty$ as $\alpha \rightarrow 0$, and it increases as α goes up, there exists a $\underline{\alpha} \in (0, 1)$, which depends on

exogenous variables and parameters, below which $(L_{TJ})_n^* > L_{TJ}^{**}$ and thus $Y_n^* > Y_e^*$ hold.

By contrast, $Y_n^* < Y_e^*$ holds if

$$(L_{TJ})_e^* < L_{TJ}^{**} \Leftrightarrow A_T(L_{TJ}^{**})^{\alpha-1} - \beta\omega_s \left(1 - 2n_e \frac{L_{TJ}^{**}}{N}\right) = \frac{A_M}{\alpha} - \beta\omega_s \left[1 - 2n_e \left(\frac{\alpha A_T}{A_M}\right)^{\frac{1}{1-\alpha}} \frac{1}{N}\right] < A_M \text{ from (22) and Assumption 2.} \quad (46)$$

From a similar reasoning as above, if $L_{TJ}^{**} = \left(\frac{\alpha A_T}{A_M}\right)^{\frac{1}{1-\alpha}} < \frac{N}{2n_e}$, there exists a $\bar{\alpha} \in (0, 1)$ ($> \underline{\alpha}$) above which $(L_{TJ})_e^* < L_{TJ}^{**}$ and thus $Y_n^* < Y_e^*$ hold. (Or, from (42), if parameters and exogenous variables are such that $(L_{TJ})_n^* \geq (L_{TJ})_e^* - (L_{TJ})_n^*$, then the inequality holds for any α when $\frac{(L_{TJ})_n^*}{N} \leq \frac{1}{6n_e + 2(n_e - 1)}$.)

■

Proof of Proposition 2. (i) $F_n^* < F_{Td}^*$ is from (26) and (32), $F_{Md}^* < F_e^*$ is from (20) and (35), and $F_{Td}^* < F_d^* < F_{Md}^*$ is from (28), (32), and (35) and $(L_{TJ})_{Td}^* < (L_{TJ})_d^* < (L_{TJ})_{Md}^*$ shown in (ii).

(ii) $(L_{TJ})_{Td}^* = (L_{TJ})_n^*$ and $(L_{TJ})_{Md}^* = (L_{TJ})_e^*$ are shown in Sections 3.2.2 and 3.2.3. As shown in footnote 25, the LHS of (22), the indifference condition whose solution is $(L_{TJ})_e^*$, decreases with L_{TJ} for $L_{TJ} \leq (L_{TJ})_e^*$.

Hence, given parameters and exogenous variables, $(L_{TJ})_d^* < (L_{TJ})_e^*$ holds, if the LHS of (29), the indifference condition whose solution is $(L_{TJ})_d^*$, is smaller than that of (22) at $L_{TJ} = (L_{TJ})_d^*$, which is true because

$$\beta \left\{ (\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e - 1}{n_e} + \omega_s \left[n_e \left(\frac{L_{TJ}}{N} \right)^2 - \left(1 - n_e \frac{L_{TJ}}{N} \right)^2 \right] \right\} - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] - \gamma (\sigma_N - \sigma_{-N}) \leq -\beta \omega_s \left(1 - 2n_e \frac{L_{TJ}}{N} \right) \quad (47)$$

$$\Leftrightarrow \beta \left[(\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] \leq \gamma (\sigma_N - \sigma_{-N}), \quad (48)$$

where the inequality holds from (87) in the proof of Proposition A2.

As shown in footnote 26, the shape of the LHS of (27), the indifference condition whose solution is $(L_{TJ})_n^*$, is similar to that of (22). Hence, $(L_{TJ})_d^* > (L_{TJ})_n^*$ holds if the LHS of (29) is greater than that of (27) at $L_{TJ} = (L_{TJ})_d^*$, which is true because

$$\beta \left\{ (\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e - 1}{n_e} + \omega_s \left[n_e \left(\frac{L_{TJ}}{N} \right)^2 - \left(1 - n_e \frac{L_{TJ}}{N} \right)^2 \right] \right\} - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] - \gamma (\sigma_N - \sigma_{-N}) \geq -\beta \omega_s \left(1 - \frac{2L_{TJ}}{N} \right) \quad (49)$$

$$\Leftrightarrow \beta \left[(\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\frac{\delta V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta-1}} - \left(\frac{\delta V}{F_d(L_{TJ})} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] \geq \gamma (\sigma_N - \sigma_{-N}), \quad (50)$$

where the inequality holds from (88) in the proof of Proposition A2.

(iii) Denote total output of the private good in the equilibrium in which those in sector M (sectors TJ) identify with the nation (their ethnic group) by Y_d^* . From (1) and (2),

$$Y_n^* > Y_d^* \Leftrightarrow A_T((L_{TJ})_n^*)^\alpha - A_M(L_{TJ})_n^* > A_T((L_{TJ})_d^*)^\alpha - A_M(L_{TJ})_d^* \quad (51)$$

$$\Leftrightarrow \beta\omega_s \left(1 - 2\frac{(L_{TJ})_n^*}{N}\right) (L_{TJ})_n^* - \left(-\beta \left\{ \begin{aligned} &[\eta_0 + \eta_1 F_d((L_{TJ})_d^*)] \frac{n_e - 1}{n_e} + \omega_s \left[\begin{aligned} &n_e \left(\frac{(L_{TJ})_d^*}{N}\right)^2 - \left(1 - n_e \frac{(L_{TJ})_d^*}{N}\right)^2 \\ &-(n_e - 1) \frac{(L_{TJ})_d^*}{N} \left(2 - n_e \frac{(L_{TJ})_d^*}{N}\right) \\ &+(n_e - 1) \frac{(L_{TJ})_d^*}{N} \left(2 - n_e \frac{(L_{TJ})_d^*}{N}\right) \end{aligned} \right] \end{aligned} \right\} \right) (L_{TJ})_d^* > 0 \quad (52)$$

$$\Leftrightarrow \beta\omega_s((L_{TJ})_d^* - (L_{TJ})_n^*) \left[\frac{2}{N}((L_{TJ})_d^* + (L_{TJ})_n^*) - 1 \right] + \left(\begin{aligned} &\beta \left\{ \begin{aligned} &[\eta_0 + \eta_1 F_d((L_{TJ})_d^*)] \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{(L_{TJ})_d^*}{N} \left(2 - n_e \frac{(L_{TJ})_d^*}{N}\right) \end{aligned} \right\} \\ &-\frac{1}{\theta} \left(\frac{n_e - 1}{n_e}\right)^{\frac{\theta}{\theta - 1}} \left[\left(\frac{\delta V}{F_d((L_{TJ})_d^*)}\right)^{\frac{\theta}{\theta - 1}} - \left(\frac{\delta V}{F_d((L_{TJ})_d^*)} - \beta\eta_1\right)^{\frac{\theta}{\theta - 1}} \right] - \gamma(\sigma_N - \sigma_{-N}) \end{aligned} \right) (L_{TJ})_d^* > 0, \quad (53)$$

where the third equation is from (27) and (29).

Since $(L_{TJ})_d^* > (L_{TJ})_n^*$ from (ii) of the proposition and the second term is non-negative from the equilibrium condition (88) in the proof of Proposition A2 (i), the above inequality holds when A_M is low enough or A_T is high enough that $\frac{2}{N}((L_{TJ})_d^* + (L_{TJ})_n^*) - 1 > \frac{4}{N}(L_{TJ})_n^* - 1 \geq 0$ is true, although the condition could hold even when $(L_{TJ})_n^*$ is much lower than $\frac{N}{4}$ if the second term is large.

From (1) and (2),

$$Y_e^* < Y_d^* \Leftrightarrow A_T((L_{TJ})_e^*)^\alpha - A_M(L_{TJ})_e^* < A_T((L_{TJ})_d^*)^\alpha - A_M(L_{TJ})_d^* \quad (54)$$

$$\Leftrightarrow \beta\omega_s \left(1 - 2n_e \frac{(L_{TJ})_e^*}{N}\right) (L_{TJ})_e^* - \left(-\beta \left\{ \begin{aligned} &[\eta_0 + \eta_1 F_d((L_{TJ})_d^*)] \frac{n_e - 1}{n_e} + \omega_s \left[\begin{aligned} &n_e \left(\frac{(L_{TJ})_d^*}{N}\right)^2 - \left(1 - n_e \frac{(L_{TJ})_d^*}{N}\right)^2 \\ &+(n_e - 1)n_e \left(\frac{(L_{TJ})_d^*}{N}\right)^2 \\ &-(n_e - 1)n_e \left(\frac{(L_{TJ})_d^*}{N}\right)^2 \end{aligned} \right] \end{aligned} \right\} \right) (L_{TJ})_d^* < 0 \quad (55)$$

$$\Leftrightarrow \beta\omega_s((L_{TJ})_e^* - (L_{TJ})_d^*) \left[1 - \frac{2n_e}{N}((L_{TJ})_e^* + (L_{TJ})_d^*) \right] + \left(\begin{aligned} &\beta \left\{ \begin{aligned} &[\eta_0 + \eta_1 F_d((L_{TJ})_d^*)] \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \left(\frac{(L_{TJ})_d^*}{N}\right)^2 \end{aligned} \right\} \\ &-\frac{1}{\theta} \left(\frac{n_e - 1}{n_e}\right)^{\frac{\theta}{\theta - 1}} \left[\left(\frac{\delta V}{F_d((L_{TJ})_d^*)}\right)^{\frac{\theta}{\theta - 1}} - \left(\frac{\delta V}{F_d((L_{TJ})_d^*)} - \beta\eta_1\right)^{\frac{\theta}{\theta - 1}} \right] - \gamma(\sigma_N - \sigma_{-N}) \end{aligned} \right) (L_{TJ})_d^* < 0, \quad (56)$$

where the third equation is from (22) and (29).

Since $(L_{TJ})_d^* < (L_{TJ})_e^*$ from (ii) of the proposition and the second term is negative from the equilibrium condition (87) in the proof of Proposition A2 (i), the above inequality holds when A_M is low enough or A_T is high enough that $\frac{2n_e}{N}((L_{TJ})_d^* + (L_{TJ})_e^*) - 1 > \frac{4n_e}{N}(L_{TJ})_d^* - 1 \geq 0$ is true, although the condition could hold even when $(L_{TJ})_d^*$ is much lower than $\frac{N}{4n_e}$ if the second term is large negative.

As for the result on the relationship between α and the magnitude relation of Y_d^* and Y_n^* or Y_e^* , the corresponding proof of Proposition 1 applies since $(L_{TJ})_e^* > (L_{TJ})_d^* > (L_{TJ})_n^*$ holds (although the condition for $Y_d^* > Y_e^*$ when α is sufficiently low and the one for $Y_d^* > Y_n^*$ when α is sufficiently high could be strengthened). $Y_{Td}^* = Y_n^*$ and $Y_{Md}^* = Y_e^*$ are straightforward from $(L_{TJ})_{Td}^* = (L_{TJ})_n^*$ and $(L_{TJ})_{Md}^* = (L_{TJ})_e^*$. ■

Proof of Proposition 4. It is enough to prove that the terms on the opposite side of $\gamma(\sigma_N - \sigma_{-N})$ of all equilibrium conditions— the LHSs of (72) and (83) in the proof of Proposition A1 and of (87) and (88) in the proof of Proposition A2— increase with η_0 (that is, all dividing lines in Figure 2 shift upward with an increase in η_0). As for the homogeneous identity equilibria, since F is independent of η_0 , the result is straightforward from (72) and (83). As for the equilibrium in which those in sector T identify with their ethnic group and those in sector M identify with the nation, since F is independent of η_0 for given L_{TJ} from (28), the result is straightforward from (87) and (88). (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, (20), (26), (28), (29), (31), and (34).

(ii) It is enough to prove that the terms on the opposite side of $\gamma(\sigma_N - \sigma_{-N})$ of all equilibrium conditions— the LHSs of (72) and (83) in the proof of Proposition A1 and of (87) and (88) in the proof of Proposition A2— increase with V (that is, all dividing lines in Figure 2 shift upward with an increase in V).

[The equilibrium in which all identify with their ethnic group] The derivative of the LHS of (72) in the proof of Proposition A1 with respect to V is, from (20),

$$\begin{aligned} & \frac{1}{\theta}(V)^{-1} \left\{ \beta\eta_1 \frac{n_e-1}{n_e} F_e^* - \left(\frac{n_e-1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F_e^*} \right)^{\frac{1}{\theta-1}} - \left(\delta \frac{V}{F_e^*} - \beta\eta_1 \right)^{\frac{1}{\theta-1}} \right] \delta \frac{V}{F_e^*} \right\} \\ & > \frac{1}{\theta}(V)^{-1} \left\{ \beta\eta_1 \frac{n_e-1}{n_e} F_e^* - \left(\frac{n_e-1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left(\delta \frac{V}{F_e^*} \right)^{\frac{\theta}{\theta-1}} \right\} = \frac{1}{\theta}(V)^{-1} \frac{n_e-1}{n_e} F_e^* \left(\beta\eta_1 - \frac{1}{N} \delta \frac{V}{F_e^*} \right). \end{aligned} \quad (57)$$

For $F_e^* = \left(\delta \frac{n_e-1}{n_e} \frac{V}{F_e^*} \right)^{\frac{1}{\theta-1}} N$ and $F_n^* = \left[\frac{n_e-1}{n_e} \left(\delta \frac{V}{F_n^*} - \beta\eta_1 \right) \right]^{\frac{1}{\theta-1}} N$ not to be very similar, $\beta\eta_1$ must be of a similar order of magnitude to $\delta \frac{V}{F_n^*}$ and $\delta \frac{V}{F_e^*}$. Then, $\beta\eta_1 - \frac{1}{N} \delta \frac{V}{F_e^*} > 0$ and thus the above derivative is positive.

[The equilibrium in which all identify with the nation] Since $\frac{dF_n^*}{dV} = \frac{\frac{1}{\theta-1} \delta}{\theta-1} \frac{\delta}{\delta \frac{V}{F_n^*} - \beta\eta_1}$ and $\frac{d\left(\frac{V}{F_n^*}\right)}{dV} = \frac{\frac{1}{F_n^*} \left(\delta \frac{V}{F_n^*} - \beta\eta_1 \right)}{\frac{\theta}{\theta-1} \delta \frac{V}{F_n^*} - \beta\eta_1}$ from (26), the derivative of the LHS of (83) with respect to V is,

$$\begin{aligned}
& \frac{\frac{1}{\theta-1}\delta}{\frac{\theta}{\theta-1}\delta\frac{V}{F_n^*}-\beta\eta_1} \left\{ \beta\eta_1 \frac{n_e-1}{n_e} - \left(\frac{n_e-1}{n_e}\right)^{\frac{\theta}{\theta-1}} \left[\left(\delta\frac{V}{F_n^*}\right)^{\frac{1}{\theta-1}} - \left(\delta\frac{V}{F_n^*}-\beta\eta_1\right)^{\frac{1}{\theta-1}} \right] \frac{1}{F_n^*} \left(\delta\frac{V}{F_n^*}-\beta\eta_1\right) \right\} \\
&= \frac{\frac{1}{\theta-1}\delta}{\frac{\theta}{\theta-1}\delta\frac{V}{F_n^*}-\beta\eta_1} \frac{n_e-1}{n_e} \left\{ \beta\eta_1 - \frac{1}{N} \left[\left(\delta\frac{V}{F_n^*}\right)^{\frac{1}{\theta-1}} - \left(\delta\frac{V}{F_n^*}-\beta\eta_1\right)^{\frac{1}{\theta-1}} \right] \left(\delta\frac{V}{F_n^*}-\beta\eta_1\right)^{1-\frac{1}{\theta-1}} \right\} \\
&> \frac{\frac{1}{\theta-1}\delta}{\frac{\theta}{\theta-1}\delta\frac{V}{F_n^*}-\beta\eta_1} \frac{n_e-1}{n_e} \left(\beta\eta_1 - \frac{1}{N}\delta\frac{V}{F_n^*} \right) > 0. \tag{58}
\end{aligned}$$

[The equilibrium in which those in sector T identify with their ethnic group and those in sector M identify with the nation] From (28),

$$\frac{dF_d(L_{TJ})}{dV} = \frac{\frac{1}{\theta-1}\delta\frac{1}{F_d(L_{TJ})} \left(\frac{n_e-1}{n_e}\right)^{\frac{1}{\theta-1}} \left[\left(\delta\frac{V}{F_d(L_{TJ})}\right)^{\frac{1}{\theta-1}-1} n_e L_{TJ} + \left(\delta\frac{V}{F_d(L_{TJ})}-\beta\eta_1\right)^{\frac{1}{\theta-1}-1} (N-n_e L_{TJ}) \right]}{1 + \frac{1}{\theta-1}\delta\frac{V}{[F_d(L_{TJ})]^2} \left(\frac{n_e-1}{n_e}\right)^{\frac{1}{\theta-1}} \left[\left(\delta\frac{V}{F_d(L_{TJ})}\right)^{\frac{1}{\theta-1}-1} n_e L_{TJ} + \left(\delta\frac{V}{F_d(L_{TJ})}-\beta\eta_1\right)^{\frac{1}{\theta-1}-1} (N-n_e L_{TJ}) \right]}, \tag{59}$$

$$\frac{d\left(\frac{V}{F_d(L_{TJ})}\right)}{dV} = \frac{\frac{1}{F_d(L_{TJ})}}{1 + \frac{1}{\theta-1}\delta\frac{V}{[F_d(L_{TJ})]^2} \left(\frac{n_e-1}{n_e}\right)^{\frac{1}{\theta-1}} \left[\left(\delta\frac{V}{F_d(L_{TJ})}\right)^{\frac{1}{\theta-1}-1} n_e L_{TJ} + \left(\delta\frac{V}{F_d(L_{TJ})}-\beta\eta_1\right)^{\frac{1}{\theta-1}-1} (N-n_e L_{TJ}) \right]}. \tag{60}$$

Thus, the derivative of the LHS of (87) or (88) with respect to V is,

$$\begin{aligned}
& \frac{\frac{1}{\theta-1}\delta\frac{1}{F_d(L_{TJ})} \left(\frac{n_e-1}{n_e}\right)^{\frac{\theta}{\theta-1}} \left[\left(\delta\frac{V}{F_d(L_{TJ})}\right)^{\frac{1}{\theta-1}-1} n_e L_{TJ} + \left(\delta\frac{V}{F_d(L_{TJ})}-\beta\eta_1\right)^{\frac{1}{\theta-1}-1} (N-n_e L_{TJ}) \right]}{1 + \frac{1}{\theta-1}\delta\frac{V}{[F_d(L_{TJ})]^2} \left(\frac{n_e-1}{n_e}\right)^{\frac{1}{\theta-1}} \left[\left(\delta\frac{V}{F_d(L_{TJ})}\right)^{\frac{1}{\theta-1}-1} n_e L_{TJ} + \left(\delta\frac{V}{F_d(L_{TJ})}-\beta\eta_1\right)^{\frac{1}{\theta-1}-1} (N-n_e L_{TJ}) \right]} \\
& \times \left[\beta\eta_1 - \frac{\left(\delta\frac{V}{F_d(L_{TJ})}\right)^{\frac{1}{\theta-1}} - \left(\delta\frac{V}{F_d(L_{TJ})}-\beta\eta_1\right)^{\frac{1}{\theta-1}}}{\left(\delta\frac{V}{F_d(L_{TJ})}\right)^{\frac{1}{\theta-1}-1} n_e L_{TJ} + \left(\delta\frac{V}{F_d(L_{TJ})}-\beta\eta_1\right)^{\frac{1}{\theta-1}-1} (N-n_e L_{TJ})} \right], \tag{61}
\end{aligned}$$

where the expression inside the large square bracket is greater than

$$\beta\eta_1 - \frac{\left(\delta\frac{V}{F_d(L_{TJ})}\right)^{\frac{1}{\theta-1}}}{\left(\delta\frac{V}{F_d(L_{TJ})}\right)^{\frac{1}{\theta-1}-1} n_e L_{TJ} + \left(\delta\frac{V}{F_d(L_{TJ})}-\beta\eta_1\right)^{\frac{1}{\theta-1}-1} (N-n_e L_{TJ})} > \beta\eta_1 - \frac{1}{N}\delta\frac{V}{F_d(L_{TJ})} > 0. \tag{62}$$

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

Proof of Proposition A1. (i) The equilibrium in which all individuals identify with their ethnic group:

[Sector M] The utility of individual i of ethnic group J in sector M equals, from (17) and (18),

$$A_M - \frac{1}{\theta}(f_{i,e})^\theta + \delta\frac{F_J}{F}V - \beta\omega_s \left(\frac{L_{TJ}}{N/n_e}\right)^2. \tag{63}$$

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

$$A_M - \frac{1}{\theta} [(f_{i,n})']^\theta + \delta \frac{(F_J)'}{F'} V - \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\left(\frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\}, \quad (64)$$

where $(f_{i,n})' = \left[\delta \frac{F_{-J}}{(F')^2} V - \beta \eta_1 \frac{n_e - 1}{n_e} \right]^{\frac{1}{\theta - 1}}$, not $(f_{i,n})' = 0$, from the assumption (16), $(F_J)' = (f_{i,n})' + \left(\frac{N}{n_e} - 1 \right) f_{i,e}$, and $F' = (F_J)' + F_{-J}$.

When N 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

$$A_M - \frac{1}{\theta} (f_{i,n})^\theta + \delta \frac{F_J}{F} V - \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\left(\frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\}, \text{ where } f_{i,n} \text{ is given by (24).} \quad (65)$$

Thus, the deviation is not profitable if

$$-\frac{1}{\theta} (f_{i,e})^\theta - \beta \omega_s \left(\frac{L_{TJ}}{N/n_e} \right)^2 \geq -\frac{1}{\theta} (f_{i,n})^\theta - \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\left(\frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\} + \gamma (\sigma_N - \sigma_{-N}) \quad (66)$$

$$\Leftrightarrow \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 - (n_e^2 - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] \right\} - \frac{1}{\theta} [(f_{i,e})^\theta - (f_{i,n})^\theta] \geq \gamma (\sigma_N - \sigma_{-N}). \quad (67)$$

[Sector TJ] The utility of individual i of ethnic group J in sector TJ is, from (19) and (18),

$$A_T (L_{TJ})^{\alpha - 1} - \frac{1}{\theta} (f_{i,e})^\theta + \delta \frac{F_J}{F} V - \beta \omega_s \left(1 - \frac{L_{TJ}}{N/n_e} \right)^2 \quad (68)$$

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

$$A_T (L_{TJ})^{\alpha - 1} - \frac{1}{\theta} (f_{i,n})^\theta + \delta \frac{F_J}{F} V - \beta \left\{ \omega_e \frac{n_e - 1}{n_e} + \omega_s \left[\left(1 - \frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\} + \gamma (\sigma_N - \sigma_{-N}). \quad (69)$$

The deviation is not profitable if

$$-\frac{1}{\theta} (f_{i,e})^\theta - \beta \omega_s \left(1 - \frac{L_{TJ}}{N/n_e} \right)^2 \geq -\frac{1}{\theta} (f_{i,n})^\theta - \beta \left\{ \omega_e \frac{n_e - 1}{n_e} + \omega_s \left[\left(1 - \frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\} + \gamma (\sigma_N - \sigma_{-N}) \quad (70)$$

$$\Leftrightarrow \beta \left((\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 + (n_e - 1) \frac{L_{TJ}}{N} \left[2 - (n_e + 1) \frac{L_{TJ}}{N} \right] \right] \right) - \frac{1}{\theta} [(f_{i,e})^\theta - (f_{i,n})^\theta] \geq \gamma (\sigma_N - \sigma_{-N}). \quad (71)$$

[The equilibrium condition] From this equation and (67), if the condition for the modern sector holds, so does the one for the traditional sector. Hence, (67) is the condition for the existence of the equilibrium when L_{TJ} is the solution for the indifference condition (22). Since ethnic groups are symmetric and thus values of aggregate variables of all groups are the same, (67) becomes

$$\beta \left[(\eta_0 + \eta_1 F_e^*) \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \left[\left(\delta \frac{V}{F_e^*} \right)^{\frac{\theta}{\theta - 1}} - \left(\delta \frac{V}{F_e^*} - \beta \eta_1 \right)^{\frac{\theta}{\theta - 1}} \right] \geq \gamma (\sigma_N - \sigma_{-N}) \quad (72)$$

$$\Leftrightarrow \beta \Delta d^2 \left[F_e^*, -n_e (n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_e^*) \geq \gamma S_N, \quad (73)$$

where F_e^* is given by (20), and $\Delta d^2[\cdot]$, $\Delta c(\cdot)$, and S_N are as defined just before the proposition.

The condition holds for any $L_{TJ} \in [0, \frac{N}{n_e}]$ when $\gamma S_N \leq \beta \Delta d^2 \left[F_e^*, -\frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*)$. When $\gamma S_N \in \left(\beta \Delta d^2 \left[F_e^*, -\frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*), \beta \Delta d^2 [F_e^*, 0] - \Delta c(F_e^*) \right)$, the condition holds for $L_{TJ} \in [0, (L_{TJ})^{\dagger\dagger}]$, where $(L_{TJ})^{\dagger\dagger}$ is the solution for $\beta \Delta d^2 \left[F_e^*, -n_e (n_e - 1) \left(\frac{(L_{TJ})^{\dagger\dagger}}{N} \right)^2 \right] - \Delta c(F_e^*) = \gamma S_N$.

(ii) The equilibrium in which all individuals identify with the nation:

[Sector TJ] The utility of individual i of ethnic group J in sector TJ is, from (25) and (24),

$$A_T (L_{TJ})^{\alpha - 1} - \frac{1}{\theta} (f_{i,n})^\theta + \delta \frac{F_J}{F} V - \beta \left\{ \omega_e \frac{n_e - 1}{n_e} + \omega_s \left[\left(1 - \frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\} + \gamma (\sigma_N - \sigma_{-N}). \quad (74)$$

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

$$A_T (L_{TJ})^{\alpha - 1} - \frac{1}{\theta} \left[(f_{i,e})'' \right]^\theta + \delta \frac{(F_J)''}{F''} V - \beta \omega_s \left(1 - \frac{L_{TJ}}{N/n_e} \right)^2, \quad (75)$$

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

When N 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

$$A_T (L_{TJ})^{\alpha - 1} - \frac{1}{\theta} (f_{i,e})^\theta + \delta \frac{F_J}{F} V - \beta \omega_s \left(1 - \frac{L_{TJ}}{N/n_e} \right)^2, \text{ where } f_{i,e} \text{ is given by (18)}. \quad (76)$$

The deviation is not profitable if⁴⁷

$$-\frac{1}{\theta} (f_{i,n})^\theta - \beta \left\{ \omega_e \frac{n_e - 1}{n_e} + \omega_s \left[\left(1 - \frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\} + \gamma (\sigma_N - \sigma_{-N}) > -\frac{1}{\theta} (f_{i,e})^\theta - \beta \omega_s \left(1 - \frac{L_{TJ}}{N/n_e} \right)^2 \quad (77)$$

$$\Leftrightarrow \beta \left((\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left\{ \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 + (n_e - 1) \frac{L_{TJ}}{N} \left[2 - (n_e + 1) \frac{L_{TJ}}{N} \right] \right\} \right) - \frac{1}{\theta} \left[(f_{i,e})^\theta - (f_{i,n})^\theta \right] < \gamma (\sigma_N - \sigma_{-N}). \quad (78)$$

[Sector M] The utility of individual i of ethnic group J in sector M equals, from (23) and (24),

⁴⁷The equation must hold with strict inequality because the deviant's approximate utility is marginally smaller than the true utility.

$$A_M - \frac{1}{\theta}(f_{i,n})^\theta + \delta \frac{F_J}{F} V - \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\left(\frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\} + \gamma(\sigma_N - \sigma_{-N}). \quad (79)$$

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

$$A_M - \frac{1}{\theta}(f_{i,e})^\theta + \delta \frac{F_J}{F} V - \beta \omega_s \left(\frac{L_{TJ}}{N/n_e} \right)^2. \quad (80)$$

The deviation is not profitable if

$$-\frac{1}{\theta}(f_{i,n})^\theta - \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\left(\frac{L_{TJ}}{N} \right)^2 + \sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 \right] \right\} + \gamma(\sigma_N - \sigma_{-N}) > -\frac{1}{\theta}(f_{i,e})^\theta - \beta \omega_s \left(\frac{L_{TJ}}{N/n_e} \right)^2 \quad (81)$$

$$\Leftrightarrow \beta \left\{ (\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s \left[\sum_{K \neq J} \left(\frac{L_{TK}}{N} \right)^2 - (n_e^2 - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] \right\} - \frac{1}{\theta} [(f_{i,e})^\theta - (f_{i,n})^\theta] < \gamma(\sigma_N - \sigma_{-N}). \quad (82)$$

[The equilibrium condition] From this equation and (78), if the condition for sector TJ holds, so does the one for sector M . Hence, (78) is the condition for the existence of this equilibrium when L_{TJ} is the solution for the indifference condition (27). Since ethnic groups are symmetric, (78) becomes

$$\beta \left[(\eta_0 + \eta_1 F_n^*) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F_n^*} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F_n^*} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] < \gamma(\sigma_N - \sigma_{-N}) \quad (83)$$

$$\Leftrightarrow \beta \Delta d^2 \left[F_n^*, (n_e - 1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \Delta c(F_n^*) < \gamma S_N, \quad (84)$$

where F_n^* is the solution for (26).

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

Proof of Proposition A2. [The proof that no other heterogenous identity equilibria exist] If workers in sector M weakly prefer to identify with their ethnic group, from (67), the following must hold in a symmetric equilibrium:

$$\beta \left[(\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] \geq \gamma(\sigma_N - \sigma_{-N}). \quad (85)$$

If workers in sector TJ weakly prefer to identify with the nation, from (78), the following must hold in a symmetric equilibrium:

$$\beta \left[(\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] \leq \gamma(\sigma_N - \sigma_{-N}). \quad (86)$$

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

(i) The equilibrium in which workers in sector T identify with their ethnic group and those in

sector M identify with the nation:

[Sector M] Because workers in sector M identify with the nation, the condition for them not to deviate from the equilibrium is given by (82) as in the equilibrium in which all identify with the nation.

In the symmetric equilibrium, the equation becomes

$$\beta \left[(\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e - 1}{n_e} - \omega_s (n_e - 1) n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta - 1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta \eta_1 \right)^{\frac{\theta}{\theta - 1}} \right] < \gamma (\sigma_N - \sigma_{-N}), \quad (87)$$

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

where $F_d(L_{TJ})$ is the solution for (28) 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 workers in sector TJ identify with their ethnic group, the condition for them not to deviate from the equilibrium is given by (71) as in the equilibrium in which all identify with their group.

In the symmetric equilibrium, the equation becomes

$$\beta \left[(\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta - 1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta \eta_1 \right)^{\frac{\theta}{\theta - 1}} \right] \geq \gamma (\sigma_N - \sigma_{-N}), \quad (88)$$

$$\Leftrightarrow \beta \Delta d^2 \left[F_d(L_{TJ}), (n_e - 1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \Delta c(F_d(L_{TJ})) \geq \gamma S_N, \quad (89)$$

where the LHS increases with L_{TJ} from (28).

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

[The equilibrium condition] Hence, the equilibrium exists iff $\beta \Delta d^2 \left[F_d(L_{TJ}), -(n_e - 1) n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_d(L_{TJ})) < \gamma S_N \leq \beta \Delta d^2 \left[F_d(L_{TJ}), (n_e - 1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \Delta c(F_d(L_{TJ}))$, where $F_d(L_{TJ})$ is the solution for (28) and increases with L_{TJ} and $L_{TJ} = (L_{TJ})_d^*$ is the solution for (29).

To be more detailed, the equilibrium exists for $L_{TJ} \in [(L_{TJ})^\sharp, \frac{N}{n_e}]$ when

$$\gamma S_N \in \left(\max_{L_{TJ}} \left\{ \beta \Delta d^2 \left[F_d(L_{TJ}), -(n_e - 1) n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_d(L_{TJ})) \right\}, \beta \Delta d^2 \left[F_e^*, \frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*) \right) \text{ and for } L_{TJ} \geq \max\{(L_{TJ})^\sharp, 0\} \text{ satisfying } \beta \Delta d^2 \left[F_d(L_{TJ}), -(n_e - 1) n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_d(L_{TJ})) < \gamma S_N \text{ when } \gamma S_N \in \left(\min_{L_{TJ}} \left\{ \beta \Delta d^2 \left[F_d(L_{TJ}), -(n_e - 1) n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_d(L_{TJ})) \right\}, \max_{L_{TJ}} \left\{ \beta \Delta d^2 \left[F_d(L_{TJ}), -(n_e - 1) n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_d(L_{TJ})) \right\} \right).$$

(ii) The equilibrium in which those in sector M are divided over their identities and all in sectors TJ identify with their ethnic group:

[Sector M] As shown in Section 3.2.2, the following indifference condition for identity choices of sector M workers must hold

$$\beta \left[(\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \left[\left(\delta \frac{V}{F} \right)^{\frac{\theta}{\theta - 1}} - \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{\theta}{\theta - 1}} \right] = \gamma (\sigma_N - \sigma_{-N}), \quad (31)$$

where F satisfies

$$F = \left(\frac{n_e - 1}{n_e} \right)^{\frac{1}{\theta - 1}} \left\{ \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{1}{\theta - 1}} P_{M,n} (N - n_e L_{TJ}) + \left(\delta \frac{V}{F} \right)^{\frac{1}{\theta - 1}} [n_e L_{TJ} + (1 - P_{M,n})(N - n_e L_{TJ})] \right\}. \quad (32)$$

Given L_{TJ} , the LHS of (31) increases with F and F satisfying (32) decreases with $P_{M,n}$. Hence, F and $P_{M,n}$ satisfying both equations exist iff

$$\begin{aligned} & \beta \left[(\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{1}{\theta - 1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta - 1}} - \left(\delta \frac{V}{F_d(L_{TJ})F} - \beta \eta_1 \right)^{\frac{\theta}{\theta - 1}} \right] \\ & < \gamma (\sigma_N - \sigma_{-N}) < \beta \left[(\eta_0 + \eta_1 F_e^*) \frac{n_e - 1}{n_e} - \omega_s n_e (n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \frac{1}{\theta} \left(\frac{n_e - 1}{n_e} \right)^{\frac{1}{\theta - 1}} \left[\left(\delta \frac{V}{F_e^*} \right)^{\frac{\theta}{\theta - 1}} - \left(\delta \frac{V}{F_e^*} - \beta \eta_1 \right)^{\frac{\theta}{\theta - 1}} \right], \end{aligned} \quad (90)$$

where F_e^* is given by (20) and $F_d(L_{TJ})$ is given by (28) and increases with L_{TJ} .

The second inequality of (90) holds for any $L_{TJ} \in [0, \frac{N}{n_e}]$ when $\gamma S_N < \beta \Delta d^2 [F_e^*, -\frac{n_e - 1}{n_e}] - \Delta c(F_e^*)$, and for $L_{TJ} \in [0, (L_{TJ})^{\dagger\dagger})$ when $\gamma S_N \in [\beta \Delta d^2 [F_e^*, -\frac{n_e - 1}{n_e}] - \Delta c(F_e^*), \beta \Delta d^2 [F_e^*, 0] - \Delta c(F_e^*)]$, where $(L_{TJ})^{\dagger\dagger}$ is L_{TJ} satisfying $\beta \Delta d^2 [F_e^*, -n_e(n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2] - \Delta c(F_e^*) = \gamma S_N$.

The LHS of the first inequality is same as (29) 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 (82) as in the equilibrium in which all identify with their ethnic group. In the symmetric equilibrium, the condition becomes

$$\beta \left[(\eta_0 + \eta_1 F) \frac{n_e - 1}{n_e} + \omega_s (n_e - 1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \frac{1}{\theta} \left[\left(\delta \frac{n_e - 1}{n_e} \frac{V}{F} \right)^{\frac{\theta}{\theta - 1}} - \left(\delta \frac{n_e - 1}{n_e} \frac{V}{F} - \beta \eta_1 \frac{n_e - 1}{n_e} \right)^{\frac{\theta}{\theta - 1}} \right] \geq \gamma (\sigma_N - \sigma_{-N}), \quad (91)$$

where F is the solution for (31) and (32). When (90) and thus (31) hold, this condition holds for certain.

[The equilibrium condition] Hence, the equilibrium exists iff $\beta \Delta d^2 \left[F_d(L_{TJ}), -(n_e - 1)n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_d(L_{TJ})) < \gamma S_N < \beta \Delta d^2 \left[F_e^*, -n_e(n_e - 1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_e^*)$, where $F_d(L_{TJ})$ is the solution for (28) and increases with L_{TJ} and $L_{TJ} = (L_{TJ})_e^*$ is the solution for (22).

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

$$\begin{aligned} & \gamma S_N \in \left(\max_{L_{TJ}} \left\{ \beta \Delta d^2 \left[F_d(L_{TJ}), -(n_e - 1)n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_d(L_{TJ})) \right\}, \beta \Delta d^2 [F_e^*, 0] - \Delta c(F_e^*) \right) \text{ and for } L_{TJ} < \\ & \min \{ (L_{TJ})^{\dagger\dagger}, \frac{N}{n_e} \} \text{ satisfying } \beta \Delta d^2 \left[F_d(L_{TJ}), -(n_e - 1)n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_d(L_{TJ})) < \gamma S_N \text{ when } \gamma S_N \in \\ & \left(\min_{L_{TJ}} \left\{ \beta \Delta d^2 [F_n^*, 0] - \Delta c(F_n^*), \beta \Delta d^2 \left[F_e^*, -\frac{n_e - 1}{n_e} \right] - \Delta c(F_e^*) \right\}, \max_{L_{TJ}} \left\{ \beta \Delta d^2 \left[F_d(L_{TJ}), -(n_e - 1)n_e \left(\frac{L_{TJ}}{N} \right)^2 \right] - \Delta c(F_d(L_{TJ})) \right\} \right), \end{aligned}$$

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

(iii) The equilibrium in which those in sectors TJ are divided over their identities and all in sector M identify with the nation:

[Sector TJ] As shown in Section 3.2.3, the following indifference condition for identity choices of sector TJ workers must hold

$$\beta \left[(\eta_0 + \eta_1 F) \frac{n_e-1}{n_e} + \omega_s(n_e-1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \frac{1}{\theta} \left(\frac{n_e-1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] = \gamma (\sigma_N - \sigma_{-N}), \quad (34)$$

where F satisfies

$$F = \left(\frac{n_e-1}{n_e} \right)^{\frac{1}{\theta-1}} \left[\left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{1}{\theta-1}} [P_{TJ,n} n_e L_{TJ} + (N - n_e L_{TJ})] + \left(\delta \frac{V}{F} \right)^{\frac{1}{\theta-1}} (1 - P_{TJ,n}) n_e L_{TJ} \right]. \quad (35)$$

Given L_{TJ} , the LHS of (34) increases with F , and F satisfying (35) decreases with $P_{TJ,n}$. Hence, F and $P_{TJ,n}$ satisfying both equations exist iff

$$\begin{aligned} & \beta \left[(\eta_0 + \eta_1 F_n^*) \frac{n_e-1}{n_e} + \omega_s(n_e-1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \frac{1}{\theta} \left(\frac{n_e-1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F_n^*} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F_n^*} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] < \gamma (\sigma_N - \sigma_{-N}) \\ & < \beta \left[(\eta_0 + \eta_1 F_d(L_{TJ})) \frac{n_e-1}{n_e} + \omega_s(n_e-1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \frac{1}{\theta} \left(\frac{n_e-1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F_d(L_{TJ})} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F_d(L_{TJ})} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right], \end{aligned} \quad (92)$$

where F_n^* is given by (26) and $F_d(L_{TJ})$ is given by (28) and increases with L_{TJ} .

Thus, the first inequality of (92) holds for any $L_{TJ} \in [0, \frac{N}{n_e}]$ when $\gamma S_N > \beta\Delta d^2 \left[F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*)$, and holds for $L_{TJ} \in [0, (L_{TJ})^{\#\#}]$ when $\gamma S_N \in \left(\beta\Delta d^2 [F_n^*, 0] - \Delta c(F_n^*), \beta\Delta d^2 \left[F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) \right)$, where $(L_{TJ})^{\#\#}$ is L_{TJ} satisfying $\beta\Delta d^2 \left[F_n^*, (n_e-1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \Delta c(F_n^*) = \gamma S_N$.

The second inequality of (92) holds for any $L_{TJ} \in [0, \frac{N}{n_e}]$ when $\gamma S_N < \beta\Delta d^2 [F_n^*, 0] - \Delta c(F_n^*)$, and for $L_{TJ} \in ((L_{TJ})^{\#}, \frac{N}{n_e}]$ when $\gamma S_N \in \left[\beta\Delta d^2 [F_n^*, 0] - \Delta c(F_n^*), \beta\Delta d^2 \left[F_e^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_e^*) \right)$, where $(L_{TJ})^{\#}$ ($(L_{TJ})^{\#} < (L_{TJ})^{\#\#}$ from $F_n^* < F_d(L_{TJ})$) is L_{TJ} satisfying $\beta\Delta d^2 \left[F_d(L_{TJ}), (n_e-1) \frac{L_{TJ}}{N} \left(2 - n_e \frac{L_{TJ}}{N} \right) \right] - \Delta c(F_d(L_{TJ})) = \gamma S_N$.

[Sector M] Because workers in sector M identify with the nation, the condition for them not to deviate from the equilibrium is given by (82) as in the equilibrium in which all identify with the nation. In the symmetric equilibrium, the condition becomes

$$\beta \left[(\eta_0 + \eta_1 F) \frac{n_e-1}{n_e} - \omega_s n_e (n_e-1) \left(\frac{L_{TJ}}{N} \right)^2 \right] - \frac{1}{\theta} \left(\frac{n_e-1}{n_e} \right)^{\frac{\theta}{\theta-1}} \left[\left(\delta \frac{V}{F} \right)^{\frac{\theta}{\theta-1}} - \left(\delta \frac{V}{F} - \beta \eta_1 \right)^{\frac{\theta}{\theta-1}} \right] < \gamma (\sigma_N - \sigma_{-N}), \quad (93)$$

where F is the solution for (34) and (35). When (92) and thus (34) hold, this condition holds for certain.

[The equilibrium condition] Hence, when L_{TJ} is the solution for (27), the equilibrium exists for $L_{TJ} \in ((L_{TJ})^{\#}, (L_{TJ})^{\#\#})$ when $\gamma S_N \in \left(\beta\Delta d^2 [F_n^*, 0] - \Delta c(F_n^*), \beta\Delta d^2 \left[F_n^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_n^*) \right)$, and exists for $L_{TJ} \in ((L_{TJ})^{\#}, \frac{N}{n_e}]$ when $\gamma S_N \in \left(\beta\Delta d^2 [F_n^*, \frac{n_e-1}{n_e}] - \Delta c(F_n^*), \beta\Delta d^2 \left[F_e^*, \frac{n_e-1}{n_e} \right] - \Delta c(F_e^*) \right)$. ■