The Value of Democracy: Evidence from Road Building in Kenya

Journal:	*American Economic Review*
Manuscript ID:	AER-2013-1031.R1
Manuscript Type:	For Administrative Use Only
Keywords:	D72, H54, O55

The Value of Democracy: Evidence from Road Building in Kenya

Robin Burgess * Remi Jedwab†

Edward Miguel‡ Ameet Morjaria§ Gerard Padró i Miquel¶

July 2014‖

Abstract

Ethnic favoritism is seen as antithetical to development. This paper provides credible quantification of the extent of ethnic favoritism using data on road building in Kenyan districts across the 1963-2011 period. Guided by a model it then examines whether the transition in and out of democracy under the same president constrains or exacerbates ethnic favoritism. Across the post-independence period, we find strong evidence of ethnic favoritism: districts that share the ethnicity of the president receive twice as much expenditure on roads and have five times the length of paved roads built. This favoritism disappears during periods of democracy.

*London School of Economics; r.burgess@lse.ac.uk

†George Washington University; jedwab@gwu.edu

‡University of California, Berkeley and NBER; emiguel@econ.berkeley.edu

§Harvard Academy for International and Area Studies; ameet.morjaria@gmail.com

¶London School of Economics; g.padro@lse.ac.uk

‖We are especially grateful to the Editor, Marianne Bertrand and four anonymous referees for suggestions. We also thank Gani Aldashev, Alberto Alesina, Tim Besley, Oriana Bandiera, Bronwen Burgess, Konrad Burchardi, Denis Cogneau, Taryn Dinkelman, Ruben Durante, Eliana La Ferrara, Greg Fischer, Paul Gertler, James Habyarimana, Charles Hornsby, Asim Khwaja, Michael Kremer, Guy Michaels, Stelios Michalopoulos, Suresh Naidu, Benjamin Olken, Torsten Persson, Jean-Philippe Platteau, Daniel Posner, Jim Robinson, Gerard Roland, Ryan Sheely, Jim Snyder, Justin Valasek, Ekaterina Zhuravskaya, and seminar audiences at UC Berkeley, CEPR Development Economics Symposium, Gothenberg, Harvard, IGC Africa Growth Conference, International Economic Association Conference, LSE, NBER, NEUDC, MIT, Namur, Oxford, Princeton, PSE, Royal Economic Society Conference, Stanford, Warwick Public Economics Conference, Warwick/Princeton Political Economy Conference, WGAPE (UCLA), World Bank ABCDE and WZB Berlin for very helpful comments. This research was funded by the ESRC-DFID Poverty Alleviation Fund, the DFID funded Institutions for Pro-Poor Growth (iiG) Research Consortium, and the International Growth Centre. Morjaria thanks Survey of Kenya and Kenya National Archives for hosting his stay in Nairobi. Padró i Miquel acknowledges financial support from the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Starting Grant Agreement no. 283837. The views expressed here are those of the authors alone and do not necessarily reflect those of DFID or any of the other organizations acknowledged here.

1 Introduction

Ethnic favoritism refers to a situation where coethnics benefit from patronage and public policy decisions, and thus receive a disproportionate share of public resources, when members of their ethnic group control the government. It has been argued by historians, political scientists, and economists that this phenomenon has hampered the economic performance of many countries, particularly in Africa (Mamdani 1996, Easterly and Levine 1997, Herbst 2000, Posner 2005, Alesina and La Ferrara 2005, Miguel and Gugerty 2005, Michalopoulos and Papaioannou 2011, Alesina et al 2012). In fact, the widespread belief among citizens that ethnic favoritism is prevalent can "poison" local political culture and make the phenomenon self-sustaining (Horowitz 1985, Esman 1994, Fearon 1999, Wamwere 2003, Chandra 2004, Padró i Miquel 2007, Caselli and Coleman 2013). According to several of these accounts, ethnic favoritism emerges when weak political institutions are unable to constrain governments from discriminating among citizens. Therefore, to understand the recent political and economic performance of many African countries, it is crucially important to determine to what extent ethnic favoritism is prevalent, and whether the emergence (or in many cases, re-emergence) of democracy has helped mitigate it.

In this paper we make two contributions. First, we quantify the extent of ethnic favoritism in public resource allocation in an African country for the post-independence period. Second, we examine whether the transition into and out of democracy affects the extent of ethnic favoritism.

These issues have been difficult to address so far due to a number of factors. To begin with, it is challenging to determine which ethnic group is getting what share of public expenditure. This problem is particularly acute in Africa where government statistical agencies have been underfunded for decades, where data on the allocation of government spending is typically patchy at best and where, even when the data does exist, there is a reluctance to release disaggregated data that could allow the populace to uncover evidence of ethnic favoritism. Moreover, estimation of ethnic bias requires observing what happens with public expenditure when there are switches of the ethnic group in power. In many African countries this is difficult given the long tenures of post-independence leaders and the fact that particular ethnic groups have tended to be dominant for extended periods. Finally, to estimate the impact of institutional changes such as democratization on ethnic favoritism, one needs to observe switches between democracy and autocracy under the same leader, which is far from common.

To address these difficulties we pick an appropriate context: road building across Kenyan districts. This setting is attractive for a number of reasons. First, there is dramatic ethnic segregation across districts in Kenya, which is the result of the design of colonial era borders in the period before Kenya's independence in 1963. Each post-independence district was dominated by a single ethnic group, and this pattern remains

stable over time. Therefore we can directly assess, using road spending and construction by district, whether or not ethnic groups that shared the ethnicity of the president disproportionately benefited from roads.

Second, road expenditure can be directly measured. We have carried out extensive historical archival work to recover road expenditure data at the project level. This has enabled us to construct district level panel data on road expenditure for all 41 Kenyan districts across the entire post-independence period. In addition, we have constructed a panel of road presence in each of the 41 Kenyan districts using historical maps. We can therefore cross-check the district road expenditure data (from the road projects) with the district road construction data (from the maps). Having this level of detailed data on two independent measures of the same public good is extremely rare in low-income countries, particularly in Sub-Saharan Africa.

Third, roads are the largest single element of public expenditure in Kenya, constituting about 15% of total development expenditure over our sample period. This is three times what the Kenyan central government spends on health, education or water. Roads expenditure is centrally allocated and a highly visible form of public investment and thus a prime area for political patronage. Road building thus represents an attractive setting in which to analyze the extent of ethnic favoritism.

Fourth, the post-independence history of Kenya provides us both with switches in the ethnicity of the president, and switches into and out of democracy under the same president (see Figure 1). During our study period, we observe (i) a transition into autocracy from democracy under the first president of Kenya (Jomo Kenyatta, an ethnic Kikuyu), (ii) a transition from a Kikuyu president to a Kalenjin president (Daniel arap Moi) under an autocratic regime, (iii) a transition out of autocracy into democracy under Moi, and (iv) a democratic succession of a Kalenjin president to a Kikuyu president (Mwai Kibaki). These shifts allow us to identify the effect of political transitions on ethnic favoritism holding the identity of the leader constant.

Fifth, as is apparent in Figure 2, democratic change in Kenya mirrors the pattern seen across Sub-Saharan Africa. Kenya, like many African countries, was reasonably democratic post-independence in the 1960s, became autocratic in the 1970s and 1980s, and then returned to democracy in the 1990s and 2000s.[1] Our results for Kenya might thus provide insights into broader patterns of African economic and political development. For example, if we find that democracy has value in terms of imposing constraints on the executive (which in turn limits ethnic favoritism), then this might help explain

[1] Polity is on a -10 to +10 scale, with scores below -5 classified as autocratic. To capture transitions between autocracy and imperfect democracy which have characterized Africa's post-independence history, we classify scores of -5 and above as democracies. This involves combining anocracies (i.e. imperfect democracies) and full democracies (which have scores 5 and above) together so that a country is either autocratic (below -5) or democratic (-5 or more) at a given point in time. We use this lower cut-off as the majority of the transitions in Africa (and indeed throughout the developing world) have been from autocracy to imperfect democracy (rather than autocracy to full democracy) and we want to exploit this variation to examine whether it affects ethnic favoritism.

why economic growth has been higher in democratic (1960s, 1990s, 2000s) relative to autocratic (1970s, 1980s) periods. We return to this issue in the conclusion.

Our unique set-up therefore allows us to assess whether there is ethnic favoritism in roads investment, to quantify the magnitude of this effect, and to estimate the extent to which favoritism is affected by democratization. To help us interpret our results, we set up a model of centralized presidential public resource allocation across districts. The model shows how the degree of ethnic favoritism is determined by the constraints on executive action that characterize different political regimes.

We find striking patterns. Across the 1963 to 2011 period, Kenyan districts that share the ethnicity of the president receive twice as much expenditure on roads and over three times the length of paved roads built relative to what would be predicted by their population share. This is evidence of an extreme degree of ethnic favoritism. However, these biases are not constant. While in periods of autocracy, coethnic districts receive three times the average expenditure in roads and five times the length of paved roads, both these biases disappear during periods of democracy. Thus, the political regime is an important determinant of ethnic favoritism. The fact that we find similar results using two independent road data sets – one based on expenditure and the other on road maps – is reassuring.

We construct a counterfactual road network based on the goal of maximizing market potential. There is no evidene of ethnic favoritism in this simulated data, nor is ethnic favoritism affected by the political regime. This indicates that (i) our strong ethnic favoritism results in the actual data series are not being driven by coethnic districts just happening to have high market potential, and (ii) our democracy results are not due to some coincidence between regime transitions and a natural expansion of the road network over time. We also show that if we drop high economic potential districts (e.g., those in the former White Highlands, around Nairobi, or on major commercial corridors) then our results still hold.

A key insight from our theoretical model is that the ethnic bias parameters that we estimate can be interpreted in terms of regime-specific constraints on executive action. In this light, our empirical findings suggest that even "imperfect" democratic institutions, like those found in Kenya during the 1960s, 1990s and 2000s, have value by imposing constraints on the executive. Indeed, we show that movements in the regime-specific executive constraint parameter derived from our model (and estimated using our data) closely parallel those in the polity measure of democracy in Figure 2. In the context of the many African countries where presidential power has an ethnic base, democracy thus may translate to lessened favoritism towards coethnics as political leaders are forced to share public goods across the wider population.

Closer examination of recent Kenyan history sheds light on how the re-emergence of democracy in the 1990s changed the nature of constraints on Kenyan leaders and

altered the allocation of public resources. Democracy heralded an increase in political choice and participation as well as less repression of popular expression, including by increasingly vocal civil society groups. There was a reduction in press censorship, an explosion of private print and electronic media and a more active role of parliament and the judiciary in holding political leaders to account (Wrong 2009). These changes meant that the actions of political leaders were under much greater scrutiny, which helps to understand why ethnic favoritism was dramatically reduced during periods of democracy.

Despite its perceived centrality to economic development in Africa, the study of ethnic favoritism in public good allocation using subnational data is a relatively recent phenomena in large part due to the absence of subnational (e.g., district) panel data sets covering the period from independence to the present. Demographic and Health Surveys for Kenya (Kramon and Posner 2014) and from across Africa (Franck and Ranier 2012), which allow researchers to construct schooling and health outcomes over long periods are generating new insights into whether or not political leaders favor coethnics. Innovative use of satellite data has also enabled researchers to track regional outcomes across leadership and regime transitions (Hodler and Rachsky (forthcoming), Morjaria 2014). The literature is thus moving beyond the seminal cross-country analysis of Easterly and Levine (1997). We will turn to a discussion of how our paper complements and contributes to this fast growing literature in the conclusion.

The remainder of the paper is organized as follows. Section 2 provides a theoretical framework. Section 3 presents the historical background on roads and politics in Kenya and describes the data. Section 4 presents the methods and results. Section 5 interprets these findings in light of our model and recent Kenyan history. Section 6 links our paper to the literature on ethnic favoritism, public goods and economic development in Africa, and concludes.

2 Theoretical Framework

Consider a repeated economy populated by infinitely lived agents that discount the future at rate δ. There is a set of citizens of size 1. Citizens belong to one of two ethnic groups, $i \in \{A, B\}$, and the population share of group A is π^A. In addition to the citizens, each group also has an elite that comprises an infinitely countable set of identical potential presidents.

At any point in time, there is a president in power who belongs to either one of these groups, $j \in \{A, B\}$. The president decides on lump-sum taxes τ, common for both groups, and on the amount of public benefits such as schooling, health, civil service jobs or roads that he provides to each group.[2] Denote by η^{ij} the per capita public benefits

[2] We assume no tax discrimination for a number of reasons. First, the empirical evidence is mixed on African governments' capacity to effectively discriminate with taxation (Bates 1981, Kasara 2007), so this simplifying assumption is a useful benchmark. Moreover, τ here includes legal taxes and also indirect ways of extracting rents. The assumption of no tax discrimination is therefore equivalent to assuming

expenditure that group i receives when the president belongs to group j. The president only cares about rent extraction, which each period is simply given by

$$\pi^A \left(\tau - \eta^{Aj}\right) + \pi^B \left(\tau - \eta^{Bj}\right).$$

The citizens of group i pay taxes τ and enjoy public benefits η^{ij}, which gives them the following simple instantaneous utility:

$$R(\eta^{ij}) - \tau,$$

where $R(\cdot)$ is increasing and concave. Note that citizens here do not have any inherent preference for the ethnicity of the president, and only care about the public benefit policies that the president implements.

The president can discriminate across groups in public spending but is limited by institutional and societal constraints. Following Besley and Persson (2010, 2011) we capture these constraints on the executive in a simple way as follows:

$$\eta^{ij} \leq \theta \left(\pi^A \eta^{Aj} + \pi^B \eta^{Bj}\right) \tag{1}$$

where $\theta \in [1, \infty]$ denotes the weakness of constraints on the executive. This formulation states that per capita spending in favor of group i cannot exceed average per capita spending by more than a factor θ. If $\theta = \infty$, institutions are so weak that they do not constrain the president in any way and all spending can be targeted to one group. At the opposite extreme, $\theta = 1$ implies that no discrimination across groups is possible.

We assume that political institutions are also relatively weak, and therefore the active support of one's coethnics is necessary to stay in power.[3] As in Padró i Miquel (2007), we assume that an acting president who receives the support of his ethnic group stays in power with probability $\bar{\gamma}$. If instead coethnics refuse to support the policies of the president, such policies cannot be implemented and he is ousted from power. In this case, an open succession follows, and the new president belongs to the same ethnic group as the ousted president with probability $\underline{\gamma}$, for $1 > \bar{\gamma} \geq \underline{\gamma} > 0$.[4]

This simple model features a unique Markov Perfect Equilibrium (MPE) characterized in the following proposition, which is proven in Appendix C.

Proposition 1 *Assume* $\theta < \max\{\frac{1}{\pi^A}, \frac{1}{\pi^B}\}$. *There is a unique MPE in which*

that the cost of rent-seeking falls equally on all citizens. Nothing crucial hinges on this assumption, since Padró i Miquel (2007) obtains similar results in a model with tax discrimination.

[3] To capture a wide variety of political institutions, we do not take a strong stance on what this support means in practice. It can range from ethnic voting for the appropriate candidate to exerting violence in order to deny other ethnic groups the full exercise of their democratic rights.

[4] In this simple formulation, the weakness of transition rules can be captured by $\bar{\gamma} - \underline{\gamma}$. A large difference captures a system where the personality of the ruler is very important, as would be the case if the ruling clique can easily manipulate the political contest. If this difference is zero, there is no personality-dependent incumbency advantage. For simplicity and to save on notation, we assume that both ethnic groups are symmetric in political terms. This might, of course, not be true in reality and both $\bar{\gamma}$ and $\underline{\gamma}$ could differ across groups, capturing differences in their populations, internal structure, or security of their hold on power. Allowing for this will not change any of the results of interest (see Padró i Miquel 2007).

1. $R'(\eta^{ii}) = R'(\eta^{jj}) = \frac{1}{\theta}$.

2. *(1) is binding for presidents of both groups.*

3. *Coethnics are indifferent between supporting and rejecting the president.*

In this model, the optimal level of public benefits spending is such that $R'(\eta) = 1$ for both groups.[5] In contrast, point 1 of the proposition says that presidents oversupply their group with public benefits and point 2 says that they only provide the other group as much as they are required by the constraints on the executive. Therefore in this model there definitely is ethnic bias in public good allocation. However, this does not mean that coethnics are much better off, since point 3 notes that the president pushes his own coethnics down to their reservation level.

To build intuition for this result, first note that the president needs the support of his group at the same time that he wants to raise τ as high as possible in order to maximize rent extraction. But for coethnic support he only needs to ensure that his group is indifferent between being under his rule or being ruled by a president from the other group. Therefore in equilibrium he can impose high taxes on everyone and partly compensate his coethnics with public benefits. This keeps his supporters indifferent since the fact that he is expropriating from them is compensated by the fear that a president from the other group would steal even more, which is true in equilibrium. Meanwhile the other group is stuck with high taxes and little public expenditure. As a result, coethnics fare better than the other group, but both groups fare much worse than under an efficient government that supplies the optimal amount of benefits and does not appropriate rents.

This rent extraction strategy hinges on the ability to discriminate. As constraints on the executive become tighter, the president is forced to provide more benefits to the other group. The more benefits he is forced to provide to the other group (i.e., the smaller θ is), the less he can appropriate and hence the weaker the incentives to manipulate public good provision to his advantage. For this reason, ethnic bias is increasing in θ.

This simple framework shows that ethnic favoritism can arise when institutions are weak, even when leaders do not value the welfare of coethnics above that of non-coethnics. It also implies that constraints on the executive are binding, which helps interpret empirical estimates. This is a general illustrative framework and is therefore not specific to any particular public good or country. The reality of Kenyan politics is, of course, more complicated than the model. For instance, while only two ethnic groups have had presidents in post-independence Kenya (Kikuyu and Kalenjin), they and several other ethnic groups were engaged in complicated coalition dynamics throughout the period that we analyze. However, we show in the results section that the bias in road construction is only tied to coethnicity with the president (and to a lesser extent the vice president), and not to coethnicity with other cabinet ministers, suggesting the focus

[5] This is because lump-sum taxes ensure that the marginal cost of public funds is 1.

on executive power is appropriate. Similarly, we show that while we observe cabinets including ministers from multiple ethnic groups, we do not observe any evidence that cabinets become more ethnically representative under democracy. These facts suggest that coalition politics are not likely to be the leading driver of ethnic favoritism under democracy or dictatorship. The model also abstracts from electoral politics which undoubtedly changed with the advent of multiparty democracy. As in the case of cabinet formation, we make this simplification because we show that there is no evidence of greater road expenditure targeting "swing voter" districts under democracy.

Therefore our simple model with only two groups and democracy working through constraints on the executive is useful for guiding our empirical analysis. Indeed, in section 5 we explicitly link our empirical findings to the θ parameter in the model and show that our interpretation of the effect of democracy in terms of constraints on the executive is consistent with the recent evolution of Kenyan political institutions.

3 Background and Data

3.1 Districts and Ethnicity in Kenya

Kenya's population comprises a mix of more than forty ethnic groups. According to the population census conducted immediately prior to independence (1962), Kenya's main ethnic groups were the Kikuyu (18.8%), Luo (13.4%), Luhya (12.7%), Kalenjin (10.8%) and Kamba (10.5%). The shares of these main ethnic groups have remained stable since then despite the fact that the national population has increased nearly fivefold (see Appendix Table A1, Panel A).

These ethnic groups predate the British but boundaries between them were often not well delineated and centralized political structures based on ethnic lines were largely absent (Sheriff 1985). Authority was typically personal and local, often a function of lineage, age and wealth and not of ethnic allegiance (Mamdani 1996, Herbst 2000).

This situation changed when the British imposed a provincial administration model in the early 20th century which divided the country into provinces and districts. In drawing district borders, the views of local African chiefs and notables were increasingly sought via boundary commissions. As shown in Appendix Figure A1, district boundaries in 1909 bear little relation to ethnic boundaries at independence. The alignment of interests between the British and local chiefs – both of whom preferred greater district ethnic homogeneity as a means of facilitating governance – however, meant that by 1933 district borders were drawn in a way so that each district typically contained a dominant ethnicity. By independence in 1963 district and ethnic boundaries tightly coincide – 38 out of 41 districts in Kenya had a single ethnic group constituting more than 50% of the population, and this remains the case up until the present. The only districts that were not dominated by a single ethnic group were Nairobi, Mombasa and Trans Nzoia.[6]

[6]Nairobi and Mombasa were (and are) the two largest cities in Kenya and Trans Nzoia is highly

In our analysis, we use the 1963 district boundaries. Districts in Kenya, in effect, serve as stable ethnic markers thus allowing us to precisely assign expenditures or road length to particular ethnic groups. This in turn enables us to establish whether districts that share the ethnicity of a given president receive more road investment and also to establish whether this bias differs across democratic versus autocratic periods.

3.2 Ethnic Politics in Kenya

African political parties were sanctioned at the Lancaster House Conference in 1960. In May of that year, the Kenya Africa National Union (KANU) was formed and led by Jomo Kenyatta (an ethnic Kikuyu). Soon after, driven by the fear of Kikuyu and Luo domination, the Kenya Africa Democratic Union (KADU) was formed. KADU was composed largely by members of numerically smaller ethnic groups, and led by Daniel arap Moi (an ethnic Kalenjin). These parties contested the first post-independence election of 1963. KANU won the election convincingly and in less than two years, all KADU MPs had joined KANU, resulting in the temporary end of opposition representation in Parliament.

In the mid-1960s, however, several members of KANU defected to a new left-leaning Luo-led party, the Kenya People's Union (KPU), which opposed the perceived growing conservatism and pro-western orientation of Kenyatta and the KANU leadership. The anti-communist logic of the Cold War, however, meant that the KPU was banned in 1969, ostensibly on national security grounds. This banning institutionalized the single-party autocracy and is reflected in a sharp fall in Kenya's polity score (Figure 2).

Kenyatta died unexpectedly of natural causes in 1978 and Moi, the sitting vice president, took power as specified in the constitution. Moi continued in the footsteps of Kenyatta and further consolidated the one-party state. Following an attempted coup in 1982 led by Kikuyu officers, he switched from a Kikuyu-Kalenjin coalition to an alliance between Kalenjins, Luhyas and numerically smaller groups, similar to the KADU alliance he had once led. The heads of parastatal enterprises, the military, police and the security apparatus were rapidly replaced with Moi's Kalenjin loyalists (Widner 1992).

The early 1990s saw an increase in both internal and external pressures for African leaders to introduce democracy, with the end of the Cold War being a catalyst (Barkan 1994). The suspension of overseas development assistance from the Paris Group of Donors forced Moi to legalize opposition parties, and Kenya held multiparty elections in 1992 for the first time since the 1960s. However, while Moi had amended the constitution to allow for multiparty competition, in parallel he had also successfully consolidated the strength of the Office of the President. His abuse of the state machinery and widespread vote fraud, together with the inability of the opposition to coordinate on a single candidate, handed Moi victory in both the 1992 and the 1997 multiparty elections.

urbanized. Economic opportunities in these agglomerations attracted a diverse group of migrants.

The return to democracy is widely accepted to have brought significant changes in the nature of Kenyan politics and civil society. The emergence of a freer press, including private ownership of media, the growth of civil society organizations and of parliamentary accountability committees, as well as a reduction in blatant human rights abuses by the security apparatus, were all arguably triggered by the emergence of political competition. These trends are not unique to Kenya, as illustrated by the Africa-wide changes in polity scores seen in Figure 2. The process put in motion by these civil society changes helped make possible the relatively free national election of 2002, which was won by the opposition for the first time, with Mwai Kibaki, an ethnic Kikuyu, becoming president, marking the country's first democratic transition of power. Moi himself did not run for president in the 2002 elections, adhering to the constitutional provision barring a third term in office. Kenya's emerging democracy has been tested since 2002 most notably in the 2007 elections.[7]

3.3 Road Investment Data in Kenya

Road building is the single largest development expenditure item in Kenya's Annual Development Budget.[8] Over the period of study, 1963-2011, road spending on average represents 15.2% of the total central government's development budget, compared to figures of 5.5%, 5.7% and 6.5% for expenditures in education, health and water, respectively.

Unlike these other public goods, which derive significant funding from local communities (in the form of *harambee* funding), investments in roads are almost entirely centrally funded and controlled. The expense and visibility of roads projects has implied that the Office of the President exercises strict oversight over road investment decisions. Requests for road projects are fed into the Ministry of Public Works by provincial and district commissioners who are nominated by (and hence loyal to) the president.[9] The Office of the President then coordinates national road funding decisions with the Ministry of Finance.

The limited availability of long-run subnational data on public goods in Kenya (and other African countries) has meant that we have had to devote considerable time and effort in constructing two measures of road investment, one based on expenditure and the other based on maps.[10] The necessary data to construct similar district-year panels for the 1963-2011 period for other public goods such as health and education do not

[7] Exit polls in 2007 suggested that Raila Odinga (an ethnic Luo) had defeated Kibaki but the electoral commission granted Kibaki victory, leading to claims of electoral fraud and widespread and intense ethnic violence.

[8] Kenya's Total Annual Budget in our study period is composed of the Development Budget and Recurrent Expenditure. Unfortunately, Recurrent Expenditure is only reported as national aggregates and thus cannot be used for district-level analysis.

[9] There was disproportionate representation from the president's ethnic group in the share of both provincial and district commissioners in the 1980s (Barkan and Chege 1989).

[10] See Appendix A for details on construction of the two road investment data series Appendix Table A2 for summary statistics on the main variables.

exist for Kenya.[11]

Our main measure of road investment is expenditures on new roads annually by district during 1963-2011, obtained from the development estimates of Kenya (see Appendix A). These contain road project level data that details the expenditure on a comprehensive list of individual road projects on an annual basis (i.e., a paved road from location A to location B through location C, at total cost X). When a road project spans locations in more than one district, we use geographic information system (GIS) data to understand the layout of the road project and quantify the relative numbers of kilometers in each district. We then decompose expenditure across the relevant districts assuming an equal distribution of costs along the construction of the total length of the road.

A convenient feature of roads is that they are easy to observe on the ground. Our second measure of road investment comes from Michelin maps, which capture the actual physical extent of paved roads. Paved roads account for the majority of road expenditures, and their spread can be reliably tracked across our period. As these maps are made by French engineers in Paris assisted by Michelin offices (mainly gas stations and tire outlets) throughout Africa, they are an independent non-governmental source of data on road investment. This data should therefore not be affected by the concern that road spending, as reported by the government, might not be accurately recorded. It is simply a measure of the physical manifestation of paved roads. Digitization of the Michelin maps thus provides us with an independent check on whether there is ethnic favoritism in road building and whether such favoritism is affected by democracy.

The limitation of this second source of data is that these maps are only published in certain years. In particular, maps were produced for the following years: 1964, 1967, 1969, 1972, 1974, 1979, 1981, 1984, 1987, 1992 and 2002. To construct our GIS data set, we use a recent GIS layer containing contemporary paved roads and then use the 11 Michelin maps in order to recreate the evolution of the paved road network backwards in time. Consistency of paved road legend labels across maps implies that we can create a district-map year panel dataset of the length of paved roads (measured in kilometers) by splicing the historical road maps with the 1963 district boundaries.

To assess what the paved road network might have looked like in the absence of ethnic favoritism, we also construct a counterfactual paved road network for the 1964-2002 period. To do this, we take the 42 urban settlements that existed in Kenya in 1962 and the 7 urban settlements near borders in neighboring countries.[12] There are 1155 potential bilateral connections between these settlements. We use the first post-independence Michelin map in 1964 to identify pre-existing bilateral connections, and we then rank the remaining settlement pairs according to their market potential, namely, the sum of the populations of the settlement pair divided by Euclidian distance between

[11] For public goods other than roads, the main issue is that we cannot disaggregate expenditures by district across the 1963-2011 period.

[12] Towns and cities with populations greater than equal to 2000 (see Appendix Figure A6).

them (Fujita, Krugman and Venables 1999). This is a commonly used metric in economic geography and trade, and is also employed by transport planners in deciding where to place roads. We then determine the total kilometers of paved road actually built between 1964 and 1967 and allocate them according to the market potential ranking of settlement pairs. We repeat this process for each map period, until we exhaust the total length of paved roads that were actually constructed in Kenya between 1964 and 2002.[13] Therefore our counterfactual network has the same length of paved roads as the real network. This counterfactual thus tells us where and when paved roads would have been built if a social planner was maximizing market potential, based on information available at independence.

Actual and counterfactual paved road networks are presented in Figure 3 for the map years which coincide with political and leadership transitions – 1969, 1979, 1992, and 2002.[14] Panel A portrays the actual paved road series and provides some useful first insights into ethnic favoritism in road construction. Between 1979 and 1992 (the Moi autocracy period), the paved road network visibly expands into Kalenjin districts whereas the road network in Kikuyu districts remains largely frozen. Then between 1992 and 2002 (the Moi democracy period) paved roads expand more evenly across the country including into districts dominated by tribes other than Kalenjins and Kikuyus. Panel B exhibits the counterfactual road series and shows a very different pattern. Roads are less concentrated in Kikuyu and Kalenjin districts and display a more pronounced "hub and spoke" pattern, whereby Nairobi and other major urban centers are connected to a wider range of towns and cities, including many that are in districts that never share the ethnicity of the president. Comparing the counterfactual and actual series we see that there is much more intensive road construction around the coastal port of Mombasa and in the non-Kikuyu and non-Kalenjin hinterlands.

In Appendix Table A3, we list the top 20 and bottom 20 bilateral road connections based on market potential. Is is clear that the top ranked pairs connect large cities to nearby settlements, many of which are not in Kikuyu or Kalenjin areas. The net result is a counterfactual road network that is much more dispersed across the country and which connects more urban centers. In Figure 3 we see that in 2002, after a decade of democracy, actual and counterfactual road networks resemble one another much more closely relative to the autocratic period 1979 and 1992 maps. Yet despite more equal treatment of different ethnic groups under democracy, non-Kikuyu and non-Kalenjin districts are never fully compensated for the lack of road investment in autocratic periods (see the 2002 maps in Figure 3).

[13] To take account of topography, we assume paved roads are constructed along the shortest, unpaved connection existing between settlements in 1964.

[14] In addition to this road length counterfactual, we also construct an analogous road expenditure counterfactual series. In Appendix Figure A2 we see that the market potential expenditure counterfactual maps for 1969, 1979, 1992, and 2002 look very similar to the market potential paved road counterfactual maps in Figure 3 (refer to Appendix B for detail on the construction of the counterfactual series).

To aid interpretation, for each of our measures of district-level road investment – the annual expenditure series, the actual paved road series, and the counterfactual paved road series – we normalize the share of national road investment that a district receives relative to the population share of that district. The main outcome variable in our empirical analysis is the share of road expenditure received by a district (out of the total national road development budget that year) divided by the population share of the district in the national population (in 1962).[15] This statistic has a natural interpretation: a value of one implies that a district received road spending that is exactly proportional to its population.[16] Values greater than (less than) one denote spending that is above (below) the national per capita average. Specifically, a value of two for this measure denotes a district that is receiving twice as much road spending as the national per capita average.

For both the actual and counterfactual paved road series, we construct a parallel measure for paved road construction (in km) per capita by district, using a measure of paved road length per capita in the district divided by average paved road length per capita nationally, as an alternative district road outcome. This measure has the same interpretation, with one denoting road construction on par with the national average, and values greater than one denoting additional construction. This data allows us to perform two distinct tests. First, we can examine whether coethnic districts get more road investment relative to their national population share and whether this varies across autocracy and democracy. Second, we assess whether our counterfactual road network, which was built to maximize market potential, shows any evidence of ethnic favoritism both across the 1963-2011 period and within and outside autocratic periods.

4 Methods and Results

4.1 Methods

We seek to estimate the relationship between the ethnicity of the president and public expenditures in districts demographically dominated by his coethnics. In the period under examination, we have Kikuyu presidents (1963-1978 and 2003-2011) and a Kalenjin president (1979-2002). There are seven districts dominated by Kikuyus and six dominated by Kalenjins, out of 41 in total. We present our results using two approaches, a graphical approach and a regression approach.

[15] If road spending in district d and year t is denoted by EXP_{dt} and district population in 1962 by $POP_{d,1962}$, while total national road spending is EXP_t and national population in 1962 is POP_{1962}, then the main road spending measure can be expressed as:

$$road_{dt} = \frac{\left(\frac{EXP_{dt}}{EXP_t}\right)}{\left(\frac{POP_{d,1962}}{POP_{1962}}\right)} = \frac{\left(\frac{EXP_{dt}}{POP_{d,1962}}\right)}{\left(\frac{EXP_t}{POP_{1962}}\right)}$$

[16] This empirical benchmark lines up with our theoretical model where the optimal path of public expenditure equalizes expenditures per capita across districts.

In our first approach, we graphically examine how the ratio of a district's share of road spending or road construction relative to its population share (i.e., $road_{dt}$) varies during the post-independence period. We divide districts in two ways. First by whether or not, in a given year, the majority ethnic group in a district is the same as that of the president. This allows us to visually assess whether districts that are coethnic with the president receive a higher share of spending on roads relative to their share in the national population. We are particularly interested in analyzing whether this bias is more or less pronounced in democratic periods relative to autocratic periods. Second, we examine the evolution of districts that are dominated by Kikuyus and Kalenjins. Since all Kenyan presidents have been either ethnic Kikuyu or Kalenjin, this allows us to examine what happens to road spending in districts when they shift in and out of being coethnic with the president. A focus here again is on whether being coethnic during autocratic periods results in districts attracting a higher share of road resources relative to democratic periods. This comparison is of particular interest as the transition from democracy to autocracy in 1969 took place under the same president (Kenyatta), as did the transition from autocracy to democracy in 1992 (Moi).

In the regression approach, our main estimating equation takes the following form:

$$\begin{aligned} road_{dt} &= \gamma_d + \alpha_t + \beta(coethnic\ district_{dt}) \\ &\quad +\delta(coethnic\ district_{dt} \times democracy_t) + \theta(X_{d1963} \times [t - 1963]) + u_{dt} \end{aligned}$$

where the dependent variable is the road spending or road construction measure for year t and district d as described above.[17] To capture coethnicity with the president, we use an indicator variable ($coethnic\ district_{dt}$) that takes a value of one for districts where at least 50% of the population has the same ethnic affiliation as the serving president. The $democracy_t$ term is an indicator variable which takes a value of one during periods of multiparty democracy (1963-1969 and 2003-2011).[18] X_{d1963} is a vector of baseline demographic, economic and geographic variables all obtained in the early to mid 1960s that might affect road spending and construction. We interact these initial conditions with linear time trends [t-1963] to allow their impact to vary over time. This allows us to control for a wide range of factors that might influence where road spending or road construction takes place. The regression also controls for district fixed effects (γ_d) and year fixed effects (α_t), and standard errors are clustered at the district level.

4.2 Graphical Analysis

The first results are presented in Figure 4. We plot the average $road_{dt}$ measure for districts that are coethnic with the president in year t and those that are not. The solid

[17] For both spending and construction we have 41 districts as defined by the 1963 district boundaries. For spending we have annual data for 49 years and hence our sample is 2009 observations. For paved road construction there are 11 Michelin maps between 1963 and 2002 and hence 410 observations as we use the change in paved road length between map periods.

[18] We define democratic years as those when the constitution of Kenya allowed multiple parties to contest elections.

vertical lines, in 1969 and 1992, capture regime transitions away from democracy and back to democracy, respectively. The broken vertical lines, in 1979 and 2002, capture presidential transitions. Two interesting patterns emerge. The first is that during periods of autocracy (the 1970s and 1980s) the ratio of district share of road expenditures to district share of population is always above one for coethnic districts and below one for non-coethnic districts, which is strongly indicative of ethnic favoritism. The second is that during periods of democracy (the 1960s, 1990s and 2000s) the ratio is consistently lower and tends to be near 1 on average for both types of districts, implying little or no favoritism.

Three transitions in Figure 4 are particularly noteworthy. The first is the rapid post-1969 rise of average $road_{dt}$ from 1 to above 2. Even with the same president in power (Kenyatta), the switch from democracy to autocracy leads to road spending more than doubling in coethnic districts over the course of a few years. The second is that this favoritism is maintained and intensified after 1979 (when Moi, an ethnic Kalenjin takes power), despite the fact that the set of districts that are coethnic with the president is now completely distinct from those pre-1979. The third is that when democracy returns in 1992 the $road_{dt}$ measure gradually falls from above 2 to around 1 even though the same president (Moi) is in place. Democracy clearly appears to have value in terms of spreading the single biggest component of Kenyan public development expenditures more evenly across districts.

As noted above, only two ethnic groups, Kikuyus and Kalenjins, produced presidents during the study period. Figure 5 categorizes districts by whether the majority of the district population is Kikuyu, Kalenjin or from another ethnic group. Kikuyu districts receive road spending in line with their population share during the early democratic period. Following the banning of opposition political parties in 1969, road spending concentrates in these districts, rising to more than double that predicted by population share. This trend of favoring Kikuyu districts ends when the Kikuyu president (Kenyatta) dies in 1978. In fact, there is a striking decline in road expenditure in Kikuyu districts, and a corresponding increase in road expenditure in Kalenjin districts timed exactly after Kenyatta's death in 1978, suggesting that Moi had the authority to rapidly divert road resources to his coethnic districts. This pattern becomes even more pronounced after the failed Kikuyu-led coup attempt in 1982.

The rise in spending on Kalenjin districts is truly meteoric: $road_{dt}$ rises from around 0.5 pre-1978 to close to 3 post-1978, representing a six-fold increase in relative road spending per capita in these districts. This highly elevated $road_{dt}$ level is maintained throughout the Moi autocratic period, as the Kikuyu $road_{dt}$ falls back down towards unity. The return of democracy under Moi in 1992 appears to reduce his ability to maintain this high degree of ethnic favoritism, and the Kalenjin district $road_{dt}$ measure drifts back down towards 1 as democracy gradually strengthens. Diminished favoritism

for districts that are coethnic with the president during periods of democracy is also associated with greater spending for the majority of districts in Kenya that are neither majority Kikuyu nor Kalenjin. As Figure 5 demonstrates, the "other" ethnic districts line has a U-shaped pattern, being close to unity in the 1960s, then falling below unity in the 1970s and 1980s, and rising back towards unity in the 1990s and 2000s. Democracy seems to have a leveling influence in ensuring that Kenyan districts receive roads resources roughly in line with their share of population irrespective of whether or not they share the ethnicity of the president.[19]

The fact that all non-Kikuyu and non-Kalenjin ethnic groups (which constitute 70% of the population at independence) get road spending allocations well below the national average during every year of Kenya's 23 years of autocratic rule, and at best achieve parity during democratic periods, is strongly indicative of misallocation in road investment.

4.3 Regression Analysis

In Table 1, we move beyond the graphical analysis and employ the regression framework specified above. Panel A of Table 1, column 1 confirms that there is strong evidence of ethnic favoritism in Kenya over the whole study period. The coefficient estimate of 0.97 in this specification implies that, on average, districts that are coethnic with the president receive roughly double the amount of roads investment relative to their share in the population. Given that roads account for approximately one sixth of all central government development spending, this represents a highly consequential degree of ethnic bias.

This central result remains robust when we sequentially add a set of factors that might influence road investment patterns. These include controls for demography (district population, area, urbanization rate – column 2), economic activity (district total earnings and employment in the formal sector, value of cash crop production for export – column 3), economic geography (being on the main Mombasa-Nairobi-Kampala corridor, bordering another country, distance to Nairobi – column 4). These controls, which are either time invariant or are measured at the start of the study period, are interacted with linear time trends to allow their effects to grow over time. Our preferred specification is that in column 4 which includes all these controls interacted with time trends. This helps to reassure us that the ethnic favoritism result is not being driven by the influence of these factors. In column 5 we observe that the result is even robust to including district specific time trends. Regardless of econometric specification, the central result that coethnic districts, on average, receive twice the level of road expenditure between

[19] In Appendix Figure A3, we break out the Kamba-Luhya-Luo ethnic groups, the three other largest ethnic groups in Kenya, from the numerically smaller groups. Both goupings exhibit the same U-shaped pattern as in Figure 5, suggesting that larger ethnic groups do not have more clout in attracting road investment and that what matters most is being coethnic with the president.

1963 and 2011 is highly robust.

We next test if ethnic favoritism is affected by whether a national democratic or autocratic regime is in place. Panel B of Table 1, column 1 indicates that ethnic favoritism in road spending falls significantly during democratic periods. Indeed, an F test indicates that there is no significant evidence of ethnic favoritism within periods of democracy in Kenya (p-value = 0.31). This is the second main result of the paper. Democracy limits the ability of the president to favor coethnics, in effect forcing him to share public resources more evenly across the population. This is equivalent to a drop in θ in our theoretical model towards unity. That even imperfect forms of democracy, such as that experienced in Kenya in the 1960s and again post-1992, can reduce ethnic favoritism in this way is a striking finding.

In the remaining columns of Panel B, we see that this second result is again robust to sequentially adding in controls for demography (column 2), economic activity (column 3) and economic geography (column 4), and to inclusion of district specific time trends (column 5). Across all columns, the F-test indicates that we cannot reject the hypothesis that ethnic favoritism in road building is absent during periods of democracy. The 1.72 coefficient on $coethnic_{dt}$ in column 4 of Panel B implies that there is almost a three fold increase in road spending in coethnic districts during autocratic periods. This can be seen in Figure 4 where our road favoritism measure rises from around 1 in the 1960s to almost 3 in the 1970s and 1980s and then falls back towards 1 in the post-1992 period. The coefficient estimates of -1.32 on the ($coethnic_{dt} \times democracy_t$) interaction in column 4 of Panel B term captures the elimination of ethnic favoritism during periods of democracy.

In Table 2 we use our second $road_{dt}$ measure, the share of the length of paved roads constructed in a district relative to its population share, and reproduce the specifications in Table 1. In Panel A we see that coethnic districts receive between three to five times the kilometers of paved roads per capita relative to the national average. This central result is robust when different initial characteristics of districts interacted with time trends are included in the regression (columns 1-4) and when we include district time trends (column 5). In our preferred specification in column 4, the coefficient estimate is 3.71, implying that coethnic districts have almost five times the length of paved roads built. Ethnic favoritism as measured by paved road construction is therefore more than twice as pronounced as that measured by road expenditures. This might be because paved roads are highly visible and signal modernization and progress, and presidents may feel that investing in them may be a more effective means of securing the support of coethnics relative to investing in non-paved roads and earthen tracks.[20]

In Panel B of Table 2 we see that the tendency to favor coethnic districts with paved roads is again greatly diminished during periods of democracy. Indeed, across all

[20] Or indeed relative to other public goods which are less visible to the public.

specifications, we find that the reduction in this bias during democratic periods is such that we cannot reject the hypothesis that ethnic favoritism is absent during periods of democracy. In column 4, the coefficient estimate on $coethnic_{dt}$ is 4.26, implying that in autocratic periods more than five times the length of paved roads are built in coethnic districts relative to the national average. The coefficient estimate on ($coethnic_{dt} \times democracy_t$) of -2.38 implies that this bias is reduced in democratic periods, and indeed the F-test (p-value = 0.33) confirms that we cannot reject that there was no ethnic favoritism during these periods.

The degree to which results match up using two independently collected data sets on road expenditure (Table 1) and road building (Table 2) is reassuring. It increases our confidence in the robustness of the two key findings of this paper: (i) there is extensive favoritism towards the president's coethnics in road investment in Kenya, and (ii) this favoritism is largely eliminated during periods of democracy.

In Table 3 we run the same specification as column 4 in Tables 1 and 2 but use our counterfactual paved road data series (see Appendix B). We construct three counterfactuals, one based on connecting settlement pairs with the largest joint populations (column 1), one based on connecting settlement pairs which are closest together (column 2) and one based on connecting settlement pairs whose market potential is the highest (column 3).[21] The main focus of our analysis is column 3 of Table 3. Consistent with Figure 3 which shows that the counterfactual road network is more dispersed that the actual road network, we find no evidence in our counterfactual simulation that coethnic districts would have more kilometers of paved road than would be predicted by their population share across the 1964-2002 period (column 3, Panel A, Table 3). This helps dispel worries that our ethnic favoritism result from Panel A of Tables 1 and 2 is being driven by the fact that coethnic districts may be receiving more road investment because they had higher market potential.[22] If paved roads had been allocated to maximize market potential, then presidential coethnic districts would not have been favored relative to non-coethnic districts.

In column 3 of Panel B in Table 3, we see no evidence of ethnic favoritism in counterfactual paved road construction in either autocratic or democratic periods. This result seems intuitive as our paved road counterfactual is based solely on the population of urban settlements at independence and the distance between these urban settlement pairs. Therefore changes between democracy and autocracy (or vice versa) should not affect where paved roads are optimally built. The result is nonetheless important as it also helps to dispel concerns that the northwestern expansion of paved roads from Nairobi intially into Kikuyu districts and then into Kalenjin districts just represented

[21] Namely, settlement pairs whose sum of populations divided by the Euclidian distance between them is largest are connected first (see Appendix B and Appendix Table A3).

[22] Column 3 of Panel A of Appendix Table A4 shows the same result for the counterfactual road expenditure series.

a natural expansion of the paved road network based on market potential, which just happened to coincide with political regime changes. Indeed as Figure 3 illustrates, the natural expansion of the road network based on market potential was towards a road network that was much more dispersed across Kenya and was unaffected by demoracy. Column 3 of Panel B in Table 3 confirms that this is the case for the counterfactual district-map year paved road panel.[23]

To summarize the results of the counterfactual road construction exercise, if we compare results for actual road expenditure (Table 1), actual paved roads (Table 2) and counterfactual paved roads (Table 3), there is clear evidence (from Tables 1 and 2) that (i) political leaders in Kenya have been skewing road investment towards coethnic districts and (ii) that democracy has largely eliminated this tendency to favor coethnic districts. The absence of both these effects using counterfactual road construction patterns in Table 3 strongly suggests that ethnic favoritism has led to misallocation in actual road construction relative to the road network that would have been built if Kenyan leaders were trying to maximize market potential.

Economic activity in Kenya is concentrated along the Mombasa-Nairobi-Kampala corridor, with the densest population settlements concentrated in the area to the northwest of Nairobi, much of which has large Kikuyu and Kalenjin populations. As an additional check that our results are not driven by some spurious correlation between coethnicity and economic potential, we drop subsets of districts which credibly could have higher market potential from the analysis and assess whether our main results still hold (see Appendix Table A5). In column 1 we drop former White Highland settler districts (located predominantly to the northwest of Nairobi) that had been the focus of economic development under British rule, in column 2 we drop Nairobi and adjacent districts, in column 3 we exclude the 15 districts on the Mombasa-Nairobi-Kampala corridor, in column 4 we exclude the 9 districts on Nairobi-Kampala corridor, and in column 5 we exclude the five richest districts in 1962. In all cases, the existence of ethnic favoritism (Panel A) and its mitigation under democracy (Panel B) remains robust. This suggests that the initial concentration of road investment in Kikuyu districts around Nairobi (under Kenyatta), followed by the shift to Kalenjin district in the northwest (under Moi) and then the spread of road investments into non-coethnic districts after democracy returned in 1992 are not simply driven by roads just tracking economic potential.

It is informative to break down the results into the five leadership periods seen in Figure 1 – Kenyatta democracy, Kenyatta autocracy, Moi autocracy, Moi democracy, and Kibaki democracy. This is needed to check whether what we are observing is a general phenomena, or one related to a particular leadership regime in Kenya. For example, we would want to know whether both early (1960s) and later (1990s, 2000s)

[23] Appendix Table A4 runs the same specficiation for the counterfactual expenditure series and also finds no evidence of ethnic favoritism in either autocratic or democratic periods.

democracy were effective in mitigating ethnic favoritism. To look at this, for each of the leadership regimes shown in Figure 1, we regress our road spending favoritism index $road_{dt}$ on indicators that capture whether a district has a majority ($\geq 50\%$) Kikuyu or Kalenjin population. The comparison districts are those that do not have either of these attributes. Coefficients from each of these five separate regressions are reported in Table 4. Guided by our model we can use these coefficients to estimate regime specific measures of constraints on the executive (θ), thus enabling us to examine how these change across regime transitions and with time varying polity scores (see section 5).

The pattern of the coefficient estimates on the Kikuyu and Kalenjin indicators across periods is telling. During the Kenyatta democracy period (1963-1969), there is no significant difference between the coefficients on the Kikuyu and Kalenjin indicators (p-value = 0.70). In the Kenyatta autocracy period (1970-1978) the Kikuyu indicator becomes positive and statistically significant, and the Kikuyu-Kalenjin difference is also statistically significant (p-value = 0.01). During the Moi autocracy period (1979-1992), things flip round and now the Kalenjin indicator is positive and statistically significant, the Kikuyu indicator is not and the two are marginally significantly different (p-value = 0.08). With the transition back to democracy during the Moi democracy period (1993-2002), both indicators lose statistical significance, as does the difference between the two (p-value = 0.14) and this pattern also holds under the Kibaki democracy period (2003-2011, p-value = 0.33). The results in Table 4 indicate that there is no evidence of ethnic favoritism in either the early (1960s) or later (1990s, 2000s) democratic periods. It is during periods of autocracy that presidents blatantly favor coethnic districts in the allocation of road spending.

Appendix Tables A6 and A7 supply some additional robustness checks. In columns 1 and 2 of Panel B in Appendix Table A6, we move to a continuous measure of coethnicity based on share of population, and the two main empirical results continue to hold. In columns 3 and 4 of Panel A we normalize the road expenditure share by the district's land area share and find that our results are robust to this normalization. In columns 5 through 8 we replicate this analysis for the paved roads measure. Across both measures we find that our results from column 4 of Tables 1 and 2 are robust to these modifications in variable construction. In Table A7 we show that our results are also robust to interacting controls with year fixed effects (column 2), to including an additional control for the number of years a district has been coethnic with the president (column 3), and to correcting for spatial clustering (columns 4 and 5, Conley 1999).

4.4 Coalition Politics

Our focus has been on the impact of being coethnic with the president on road spending and paved road construction within a district, and on whether this changes under democracy. We find this makes sense given the nature of politics in many Sub-Saharan

African countries, where presidents traditionally enjoy considerable personal decision-making authority. However, it is possible that other members of the president's cabinet also influence where road investment takes place. This introduces a set of related but distinct issues pertaining to inter-ethnic coalition formation. A lessening of ethnic favoritism under democracy, for example, may not be due to changing constraints on the president alone but rather to cabinets becoming more representative, or non-coethnic groups being targeted as a means of securing votes. While a full treatment of these issues is beyond the scope of this paper, and does not feature in our theoretical model, we use our data to explore whether considering coalition politics significantly changes any of our main conclusions.

We assembled a data set that codes the ethnicity of each cabinet member for each of the thirteen central government cabinets between 1963 and 2011 (Panel B, Appendix Table A1). In an exhaustive set of regressions, we tested whether districts that are coethnic with the Public Works minister, or with ministers holding the most important cabinet portfolios (e.g., Finance, Home) receive more road spending but cannot reject the hypothesis that these effects are zero (not shown). This is further confirmation of the overriding power of presidents relative to other public officials in post-independence Kenya.

However, in column 2 of Table 5 we show the one exception: we find that districts that are coethnic with the vice president do have road expenditures significantly above the national average. A coefficient of 1.46 on the $VP\text{-}coethnic_{dt}$ measure tells us that during autocratic periods, districts receive two and a half times the average amount of road expenditure relative to their population share, a large effect. The coefficient of -1.44 on ($VP\text{-}coethnic_{dt} \times democracy$) implies that this ethnic favoritism is non-existent during periods of democracy, as is also confirmed by the *F-test* in column 2. What is also interesting in column 2 is that, during autocratic periods, districts that are coethnic with the president receive three and half times the amount of road expenditure relative to districts that are neither coethnic with the president or vice president. This finding confirms that the president has been the dominant force in allocating road spending, but also shows that the vice president is able, to a more limited extent, to skew resource allocation.[24] The fact that both these forms of favoritism dissapear during democratic periods suggest that democracy partially ties the hands of both top executives.

It is often argued that the typical way coalition politics play out in African settings is in cabinet formation. Our data set on the ethnicity of all cabinet ministers for election years between 1963 and 2011 reveals that Kenyan cabinets have been surprisingly representative, incorporating many ethnic groups beyond that of the president even during periods of autocracy (Panel B, Appendix Table A1). This is in line with what Francois et al (forthcoming) find for the post-independence cabinets of 15 African countries. When

[24] Throughout the post-independence period, the vice president was never of the same ethnicity as the president.

we regress the ethnic cabinet share divided by population ethnic share on an ethnic group indicator (which equals one if the group is coethnic with the serving president) we find that the president's group receives 65% more cabinet posts (Panel A, column 3, Table 5). This again is in line with Francois et al (forthcoming), who find that the leader's ethnic group receives a disproportionate share of cabinet posts in the countries they study. However, when we interact the group indicator with $democracy_t$ in Panel B, we find no interaction effect, indicating that the propensity to favor coethnics with cabinet positions is not attenuated during periods of democracy. In column 4 we see that the ethnic groups of both the president and vice president are favored with cabinet positions but that once again neither of these patterns is significantly affected by democracy.

These patterns are informative in at least two respects. First, there is indeed a propensity for presidents and vice presidents to "stuff" the cabinet with coethnics. Second, this tendency is not checked by the arrival of democracy, which suggests that the relationship between ethnic favoritism and democratization that we estimate is very unlikely to be driven by changes in cabinet composition.

Allocating cabinet positions to ethnic groups that are not coethnic with the president may help reduce the threat of revolution from outsiders and coups from insiders (Francois et al (forthcoming)). However, what our results indicate is that this representation does not translate into enhanced road investment in the districts that share the ethnicity of these non-coethnic ministers. It has primarily been the president who retains the power to allocate public road resources in Kenya, and democracy constrains this power without substantially changing ethnic coalition politics, at least as reflected in cabinet composition.

Another possibility is that presidents may target road investments to districts dominated by large non-coethnic groups when democracy arrives not because they are constrained in their actions, but because this may be an effective means of securing swing votes in competitive elections. In column 5 of Table 5 we see that the coefficient estimate on the Kamba-Luhya-Luo district indicator interacted with democracy is not statistically significant.[25] This implies that districts dominated by these ethnic groups, who are likely to be pivotal in elections, do not receive additional road investments relative to other non-coethnic groups when the country becomes democratic. In column 6 of Table 5 we interact our democracy measure with an indicator for non-coethnic districts where one ethnic group comprises less than 80% of the population. When democracy arrives, these relatively ethnically mixed districts (which may naturally be more competitive politically) do not receive more road investment than less mixed districts. In columns 5 and 6 of Appendix Table A8, we use the margin of victory (the winner's minus the runner up's vote share) and a party competition Herfindhal index (both from the 1992 election) interacted with democracy, and we once again find no evidence that dis-

[25] These three large ethnic groups constituted 37% of the Kenyan population in 1962 and have occupied a similar share of the population since then (Panel A, Appendix Table A1).

tricts where political competition was more intense receive more road investment when democracy returns to Kenya after 1992.

5 Interpretation

The results indicate that our theoretical framework – where two ethnic groups compete to produce presidents, the president makes all public good allocation decisions, and his ability to favor coethnics with public goods is limited by constraints on the executive – represents a useful vehicle for (i) understanding why ethnic favoritism might arise and (ii) interpreting the coefficients estimated. In this section we first use the model to derive regime-specific estimates of constraints on the executive (θ). This allows us to track changes in constraints across democratic and autocratic periods even for the same leader. We then use material from a wide variety of sources to identify what possible factors underlie changes in θ, focusing on key institutional and political reforms that have taken place since 1992.

Recall that θ captures the ability of the executive to discriminate across ethnic groups. If $\theta = 1$ then all ethnic groups receive a public good allocation equal to the average per capita allocation and ethnic favoritism is therefore impossible. If $\theta = \infty$, then the executive is unconstrained as regards the extent to which public good allocation to his ethnic group can exceed the average allocation. The fact that (1) is binding in equilibrium allows us to derive empirical estimates of θ. Specifically, our estimate β can be expressed in terms of the model as

$$\beta = \frac{\eta^{AA} - \eta^{BA}}{\pi^A \eta^{AA} + \pi^B \eta^{BA}}.$$

We can thus estimate a specific β for each regime and translate it into regime specific θ's using the fact that $\theta = 1 + \beta(1 - \pi^A)$ for each of the five regimes.[26] This enables us to trace the evolution of θ across the five periods shown in Figure 1: Kenyatta democracy, Kenyatta autocracy, Moi autocracy, Moi democracy, and Kibaki democracy. The results are presented in Figure 6. We also include the polity score for Kenya from Figure 2 in this figure.[27]

There is a remarkable correspondence between these two measures over time. The early democratic period in the 1960s was characterized by relative democratic freedoms, and essentially no evidence of ethnic favoritism towards Kenyatta's Kikuyu ethnic group, with the estimated θ near 1. However, there is a sharp increase in θ after 1970, when democracy was abandoned, with θ moving higher towards a value of 2. Polity scores move in tandem, dropping precipitously around 1970, signaling a collapse in democratic

[26] The transformation uses the fact that $\eta^{AA} = \theta\left(\pi^A \eta^{AA} + \pi^B \eta^{BA}\right)$ and $\eta^{BA} = \frac{\left[\pi^A \eta^{AA} + \pi^B \eta^{BA}\right] - \pi^A \eta^{AA}}{\pi^B}$ to generate the expression $\beta = \frac{\theta - 1}{\pi^B}$. π^A captures the population share of the ethnic group that is coethnic with the serving president. This value varies across periods as the president's ethnicity changes.

[27] Note that the θ score is presented with a reverse axis to facilitate comparison with the polity score.

freedoms, and staying low until the early 1990s. θ moves even higher during Moi's single-party rule (1979-1992), reaching 2.68, implying that the president's coethnic districts received more than two and half times more road funds on average than other groups. However, θ moves back towards 1 when democracy was restored in late 1991 and ends up nearly equal to 1, indicating that there is effectively no ethnic favoritism in the most recent period, which is the most democratic on record for post-independence Kenya. Polity scores also rise sharply in the early 1990s, signaling a return to democratic freedoms, and actually, by the 2000s, achieve levels which exceed those seen in the 1960s. Figure 6 thus indicates that, during the autocratic 1970s and 1980s, presidents are less constrained in their ability to skew road spending towards coethnic districts relative to the democratic 1960s, 1990s and 2000s. The value of democracy lies in its ability to tie the hands of presidents so that they cannot allocate public resources in a discriminatory way.

The co-movement of θ and the polity measure of democracy in Figure 6 begs the question of what underlies the changes in θ. Digging into the various components of the polity measure sheds some light into the institutional changes occurring in Kenya during its political transitions. Closer examination of Figure 6 reveals that the combined polity score decreased from 0 to -7 in the transition out of democracy during Kenyatta's leadership. Almost all sub-components of the score changed at that time: competitiveness and openness of executive recruitment worsened (there was only one party now, whose leader was chosen for life), constraints on the chief executive weakened (the president could generally bypass parliament), regulation of political participation became restrictive (participation was restricted to life members of the single-party and civil society was heavily repressed) and competitiveness of participation was eliminated (there was only one candidate for the executive seat). It is little wonder that the president under this autocratic regime felt free to allocate resources largely as he wished.

The reform of the constitution and the return to democracy in 1992 led the combined polity score to improve from -7 to -5 and up to -2 in 1997 as parties were allowed to compete and KANU's tight grip on civil society gradually loosened (this process increased scores on both regulation and competitiveness of political participation). This movement from -7 to -2 represents a very significant improvement in fundamental democratic freedoms. After the democratic presidential transition of 2002, other components of the polity score improve and push the overall score sharply higher, to around 8. Our estimated θ matches this path: from post-1992 to 2002, the estimated θ equals 1.62 but this drops to 1.00 after 2002. 1.62 represents a significant increase in constraints on the executive relative to the Moi autocratic years ($\theta = 2.68$) but falls short of the more fully constrained post-2002 setting ($\theta = 1.00$) where ethnic favoritism in road investment has largely disappeared.

We now turn to examining different factors that might underlie these changes in

θ. Information on where roads were being built in the press certainly seems to have increased after the arrival of democracy in 1992. In Appendix Figure A4 we have catalogued the number of stories pertaining to roads during the 1985 to 2010 period for the two daily, independent newspapers with the largest national circulation, the *The Daily Nation* and *The Standard*. Using a team of two Kenyan journalists (who were not informed of our research question) we counted the number of stories relating to roads in each of the daily editions of these two newspapers in 1985-2010. In Appendix Figure A4 we see that the number of stories referring to roads in *The Daily Nation* jumps abruptly about a year after the arrival of democracy, in December 1992. The same pattern is seen for *The Standard* though the rise occurs slightly earlier, in 1991. The increases in road reporting are not small: pre-1992 the number of road stories in the *The Daily Nation* is around 25 per annum rising to about twice that after 1992. The rise in road reporting in *The Standard* is more moderate, with comparable numbers being 35 and 54. The fact that the pattern is the same for two separate newspapers is reassuring and indicates that newspapers are conveying more information about road investments after democracy.[28] Working out how this information is being utilized is beyond the scope of this paper, but the fact that it is more available is an important change. These developments line up with country-level measures which try to capture the freedom of civil society institutions. A plot of the Freedom House *Freedom of the Press Index*, for example, reveals that press freedom moved from "not free" to "partly free" when the switch to multiparty politics occurred (not shown).

Broadcast media (TV and radio) on the other hand remained more firmly in the government's grip. The two TV stations *Kenya Broadcasting Company* and *Kenya Television Network* initially continued to be subject to state oversight in the post-democratic period.[29] The situation was similar for radio, a major source of information for the rural majority, which saw no independent radio licenses granted until 1996. The situation, however, improved dramatically after 1998 when state censorship of broadcast media was abolished, and by 2000 Kenya had 9 private TV stations and 19 radio stations. While state harassment has not totally disappeared, it is undeniable that mass media has become much freer since the early 1990s.

The number of non-governmental organizations (NGO's) also grew rapidly in the 1990s. While Kenyan law does not allow international donors to fund opposition political parties, they did fund governance-focused civil society organizations. Aid was also increasingly channelled through NGOs and by the late 1990s, Kenya had among the highest concentration of NGOs per capita in Sub-Saharan Africa. The churches, often in tandem with NGOs, also played a crucial role in the 1990s in giving voice to the need for impartial conduct of elections and voter registration reforms.

[28] Roads as a share of total development expenditure is similar on either side of 1992 indicating that the large rises we observe are not just a function of increases in road investment after democracy arrives.

[29] All twenty applications to start new TV stations between 1985-1995 were rejected.

A freer press and a stronger civil society, together with Western donor pressure, eventually made Moi realize that he had to accommodate demands for further openness. This brought about three key reforms in 1997, the so-called Inter-Parliamentary Parties Group (IPPG) reforms. The IPPG reforms reduced state internal security powers (e.g., preventive detention) and amended the Public Order, Broadcasting, and Societies Acts.[30] The final years of the Moi regime also saw a rise in the power of parliament, with constitutional amendments that increased its independence from the executive branch.

Institutional reforms which place greater scrutiny on the actions of the president have continued. A new constitution was ratified by voters in 2010 that altered the division of powers between the central government and newly created (and popularly elected) county governments, and consolidated a more independent judiciary. Nowadays, Kenya's increasingly well-informed, educated, and connected population is highly politically engaged. Parliamentary debates are frequently shown on national TV and discussion forums are held to allow for civil society feedback. Misguided public investments and corruption remain widespread but are more regularly brought to light by the press (Wrong 2009).

It is hardly surprising that ethnic favoritism in public resource allocation is now much more difficult to carry out than in the past. Ethnic divisions have not disappeared, and they remain highly politically salient, as tragically demonstrated in the post-election violence in 2007-08. However, freer flows of information, a vocal civil society and an independent parliament all severely curtail the ability of the executive to blatantly discriminate between different districts in choosing where to place roads projects. This is succinctly captured in our estimated $\theta = 1.00$ for the post-2002 period.

6 Conclusion

For ethnic favoritism to be a viable political strategy, the president must be able to manipulate the allocation of public expenditure with few constraints and little political cost. Ethnic favoritism and weak controls on the chief executive thus go hand in hand. As democracy becomes consolidated in many low-income countries, including many in Sub-Saharan Africa, not only does political competition become better regulated, but the constraints on executive action are also strengthened due to the scrutiny that parliament, mass media and civil society are able to exercise. In this paper, we examine this logic in detail by asking two empirical questions. First, can we detect quantitative evidence of ethnic favoritism in public resource allocation in an African country? Second, does the transition into and out of democracy under the same leader exacerbate or constrain this ethnic favoritism?

[30] The Public Order Act was amended to remove the need to obtain a license before meetings, replacing it with a need to notify the police. The Broadcasting Act was changed to provide free airtime to all parties and to promote a balanced show of opinions. The Societies Act was amended to require the registrar to respond reasonably to all requests for voter registration within 120 days.

Though many of Africa's ills have been blamed on ethnic favoritism, it has been surprisingly difficult to find concrete evidence of this behavior, mostly due to lack of data. Therefore, to address these questions we construct two new data series that directly capture public allocation decisions by the central government, one based on the geographic coding of road project data and the other on the innovative use of historical maps. We are helped in this respect by the fact that each Kenyan district is dominated by a particular ethnic group, which allows us to precisely assign expenditures or road length to ethnic groups. In answering the second question, we are helped by the fact that there have been multiple switches of power between leaders of different ethnic groups in Kenya and, within each ethnic regime, switches between democracy and autocracy.

There are two main empirical results. First, central government investments in roads have been subject to a high degree of ethnic favoritism, with districts coethnic to the president receiving three times the average expenditure in roads and five times the length of paved roads during periods of autocracy. In contrast, ethnic groups not linked to the president, which constitute the bulk of the population, receive far fewer roads across the 23 year autocratic period. Second, these biases disappear almost entirely during periods of democracy. This more equal treatment, however, is not enough to overturn the roads deficit that non-coethnic groups accumulated over autocracy.

Our result on the presence of ethnic favoritism is broadly in line with an innovative set of recent papers that use recall data on fertility and the health and schooling outcomes of children in Demographic and Health Surveys to construct panel data on infant mortality and years of schooling that span different presidential regimes. These outcome measures, which reflect the combined human capital investment decisions made by households, communities and governments, are useful complements to our direct measure of central government road investment. Kramon and Posner (2014) show that Kenyan citizens who are coethnic with the president, education minister and the health minister are more likely to attend and complete primary and secondary school. Franck and Rainier (2012) use household data for 18 African countries to show that being coethnic with the political leader leads to lower infant mortality and a higher probability of completing primary school.[31]

Our result that democracy mitigates ethnic favoritism also requires wider investigation given that autocracy has been rapidly declining both in Africa and around the world (see Figure 2 and Appendix Figure A5). Hodler and Rachsky (forthcoming) move the literature in this direction, using subnational data from across the world for 1992 to 2009 to show that the region of birth of the national political leader shows greater night light density shortly after he takes office. This effect is muted during periods of democracy. Indeed, when a country's polity score exceeds 6, birth regions of the political leader are no longer favored, which is in line with our findings for Kenya. The global

[31] Individual regressions for their 18 countries reveal positive, significant effects for these outcomes in wide range of countries suggesting that ethnic favoritism is widespread in Africa.

scope of that paper, plus the fact that light intensity captures the influence of a range of public goods as well as economic development per se, make it an interesting complement to our paper.

Linking our findings to aggregate economic outcomes represents a key priority for future research.[32] Figure 7 and Table 6 represent a first attempt in this direction. In Figure 7 we see that economic growth in Kenya and in Africa as a whole are highest during the the democratic periods (the 1960s, 1990s and 2000s) and falls towards zero during autocratic periods (1970s and 1980s). It is striking in Figure 7 that growth collapses precisely when Kenya becomes autocratic (in 1969) and then rises again when democracy returns (in 1992). Comparing Figure 2 with Figure 7, a similar picture emerges for Africa as a whole. Of course, many factors beyond ethnic favoritism might lie behind these patterns. Yet if we take the oft cited negative relationship between ethnic favoritism and economic performance seriously, then the reduction of ethnic favoritism during periods of democracy could have contributed to higher economic growth during these periods.[33]

In Table 6 we probe this idea further by extending the influential Easterly and Levine (1997) analysis. In column 1, we replicate their key result using cross-country data for the whole world from the 1960s to the 1980s, which shows ethnic fractionalization is negatively associated with economic growth. Column 2 extends the Easterly-Levine data set to the 2000s. The ethnic diversity-growth relationship is now smaller in magnitude and no longer statistically significant. This is interesting because the 1990s and 2000s were the period when many countries across the world became democratic (see Appendix Figure A5). In column 3 we test whether the association between ethnic fractionalization and economic growth varies with the presence of democracy. The results are striking: while the negative relationship Easterly and Levine (1997) uncovered still holds for autocracies, there is no association between ethnic fractionalization and economic growth in democracies. Column 4 shows that if we restrict the sample to Africa, the negative ethnic diversity-growth relationship still holds in autocracies but again is eliminated in democracies, thus paralleling our findings for Kenya.

Obviously, these cross-country results cannot necessarily be taken as causal, since democratization may be correlated with other important societal changes and is far from randomly assigned. Nonetheless, we view these patterns as useful for motivating further research. Particularly high on this research agenda is gaining a better theoretical and empirical understanding of how democracy fosters institutional changes which constrain public resource misallocation and underpin economic growth. This is as relevant for

[32] Acemoglu et al (2014) exploit the fact that transitioning to democracy or autocracy is highly correlated across countries in the same region (see Figure 2) to provide evidence that democracy has a significant positive effect on growth.

[33] A recent literature emphasizes how investments in transportation infrastructure can increase productivity and growth (see Michaels 2008, Donaldson (forthcoming)). Our results suggest that these resources were misallocated during autocracy, which may help explain why economic growth was depressed during the 1970s and 1980s in Kenya.

Kenya as it is for Myanmar and for the broad range of countries that are moving from autocracy into some form of imperfect democracy.

7 References

Acemoglu, D., S. Naidu, P. Restrepo, and J. Robinson (2014) "Democracy Does Cause Growth," Working Paper, MIT.

Alesina, A., and E. La Ferrara (2005) "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 43(3), 762-800.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2012) "Ethnic Inequality," NBER Working Paper 18512.

Barkan, J. D. and M. Chege (1989) "Decentralising the State: District focus and the Politics of Reallocation in Kenya," *The Journal of Modern African Studies*, 27(3).

Barkan, J.D. (ed.) (1994) *Beyond Capitalism Vs. Socialism in Kenya and Tanzania.* Boulder, CO: Lynne Reinner Publishers.

Bates, R.H. (1981) *Markets and States in Tropical Africa.* Berkeley: University of California Press.

Besley, T. and T. Persson (2010) "State Capacity, Conflict, and Development," *Econometrica*, 78(1), 1-34.

——— (2011) "The Logic of Political Violence," *Quarterly Journal of Economics.* 126(3), 1411-1445.

Caselli, F. and F. W. Coleman (2013) "On the Theory of Ethnic Conflict," *Journal of the European Economic Association* 11(1), 161-192.

Chandra, K. (2004) *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India.* Cambridge: Cambridge University Press.

Conley, T.G. (1999) "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 92(1), 1-45.

Donaldson, D. (forthcoming) "Railroads of the Raj: Estimating the Impact of Transportation," *American Economic Review.*

Easterly, W. and R. Levine (1997) "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, 112 (4), 1203-1250.

Esman, M. (1994) *Ethnic Politics*, Cornell University Press.

Fearon, J. (1999) "Why Ethnic Politics and "Pork" Tend to Go Together," Working Paper, Stanford University.

Franck, R. and I. Ranier (2012) "Does the Leader's Ethnicity Matter: Ethnic Favoritism, Education and Health in Sub-Saharan Africa," *American Political Science Review* 106(2), 294-325.

Francois, P., I. Rainer, and F. Trebbi (forthcoming) "How is Power Shared in Africa?," *Econometrica.*

Fujita, M., P. R. Krugman, and A. J. Venables (1999) *The Spatial Economy, Cities, Region and International Trade.* MIT Press.

Herbst, J. (2000) *States and Power in Africa.* Princeton University Press.

Hodler, R. and P. A. Raschky (forthcoming) "Regional Favoritism," *Quarterly Journal of Economics.*

Horowitz, D. L. (1985) *Ethnic Groups in Conflict.* University of California Press.

Kasara, K. (2007) "Tax Me If You Can: Ethnic Geography, Democracy, and the Taxation of Agriculture in Africa," *American Political Science Review*, 101(1), 159-172.

Kramon, E. and D. Posner (2014) "Education For All? The Political Economy of Primary Education in Kenya", Manuscript, UCLA.

Mamdani, M. (1996) *Citizen and Subject: Contemporary Africa and Legacy of Late Colonialism.* Princeton University Press.

Michaels, G. (2008) "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *Review of Economics and Statistics* 90(4), 683-701.

Michalopoulos, S. and E. Papaioannou (2011) "The Long-Run Effects of the Scramble for Africa," NBER Working Paper 17620.

Miguel, E. and M. K. Gugerty (2005) "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics*, 89(11-12), 2325-2368.

Morjaria, A. (2014) "Is Democracy Detrimental for the Environment in Developing Countries? Evidence from Kenya," Working Paper, Harvard.

Padró i Miquel, G. (2007) "The Control of Politicians in Divided Societies: The Politics of Fear," *Review of Economic Studies*, 74(4), 1259-1274.

Posner, D. (2005) *Institutions and Ethnic Politics in Africa.* New York: Cambridge University Press.

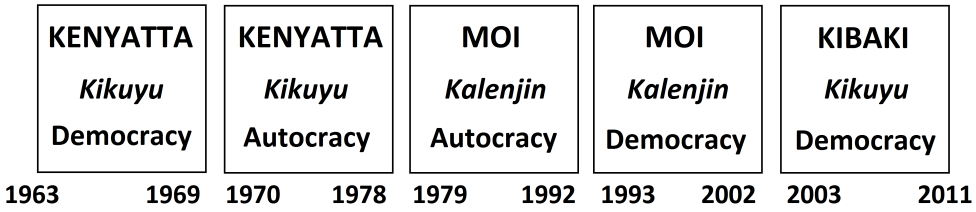
Sheriff, A. M. H. (1985) "Social Formations in Pre-Colonial Kenya", Ogot, B. A. (ed.) *Hadith 8: Kenya in the 19th Century*, Kisumu: Anyange Press.

Wamwere, K.W. (2003) *Negative Ethnicity: From Bias to Genocide.* New York: Seven Stories Press.

Widner, J. (1992) *The Rise of a Party-State in Kenya: From "Harambee!" to "Nyayo!".* Berkeley: University of California Press.

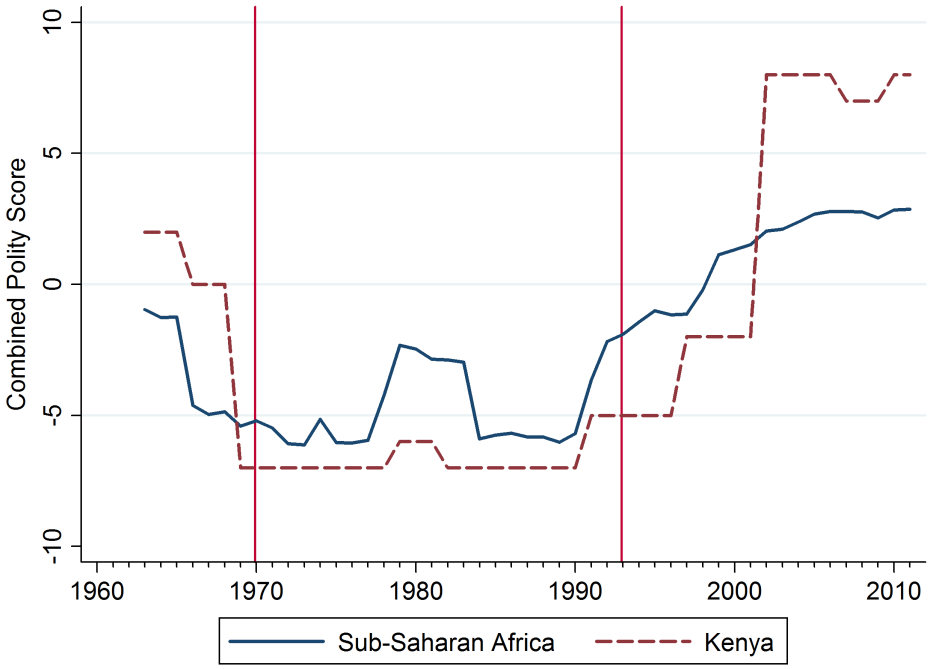
Wrong, M. (2009) *It's Our Turn to Eat: The Story of a Kenyan Whistle-Blower.* London: Fourth Estate.

Figure 1: Political and Leadership Transitions in Kenya, 1963-2011



Notes: This timeline illustrates the history of political transitions and leadership transitions in Kenya. Political transitions are as follows: December 1969 is the transition from democracy to autocracy, while December 1992 is the return of democracy. Leadership transitions: from Kenyatta (Kikuyu) to Moi (Kalenjin) in August 1978, and from Moi (Kalenjin) to Kibaki (Kikuyu) in December 2002.

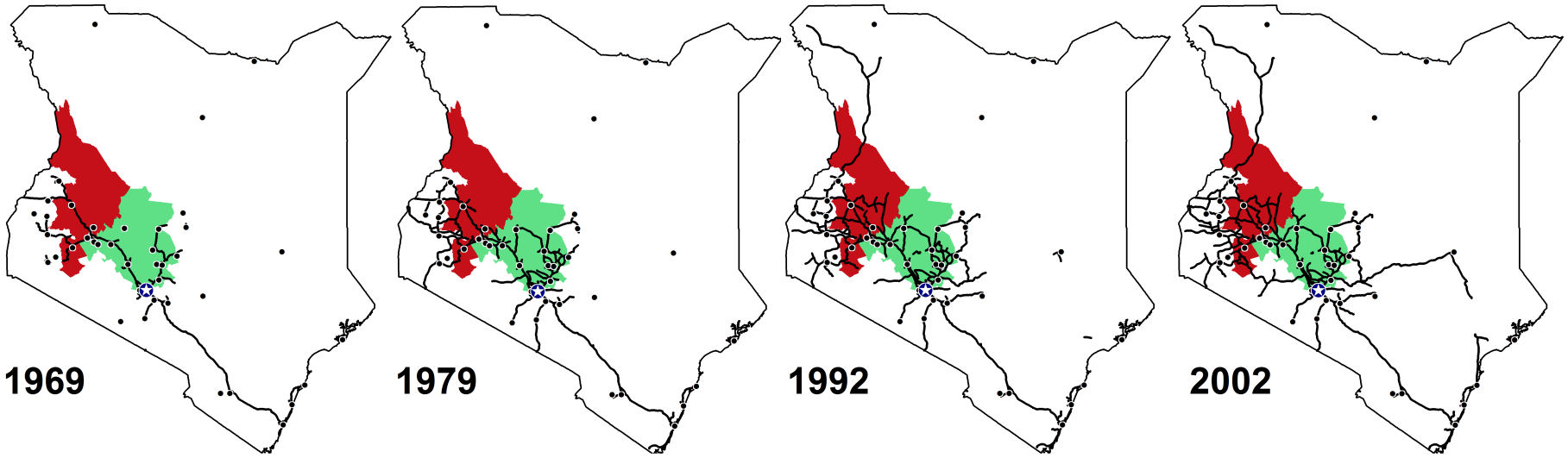
Figure 2: Evolution of Political Regimes in Sub-Saharan Africa, 1963-2011



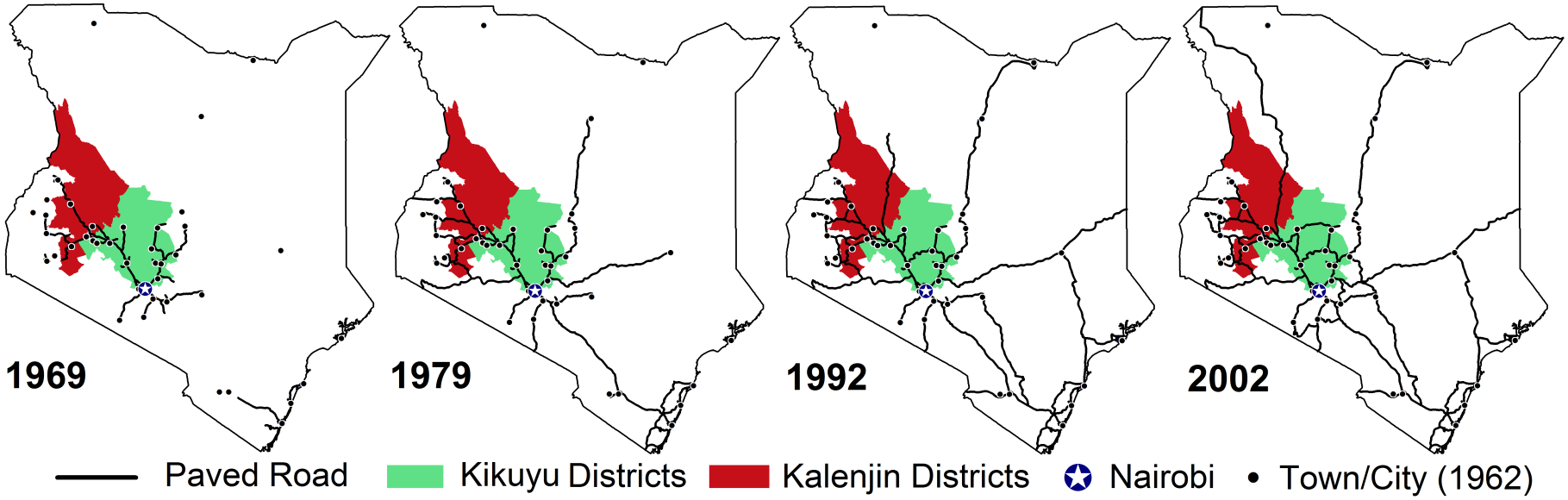
Notes: This figure plots the revised combined polity score for Kenya and the population weighted average for the rest of Sub-Saharan Africa. Polity IV defines regimes in three categories: autocracies (-10 to -6), anocracies (-5 to +5) and democracies (+6 to +10). Red vertical lines indicate regime changes in Kenya: December 1969 is the transition from democracy to autocracy, while December 1992 is the return of democracy. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Figure 3: Evolution of Kenya's Paved Road Network

Panel A: Actual Network

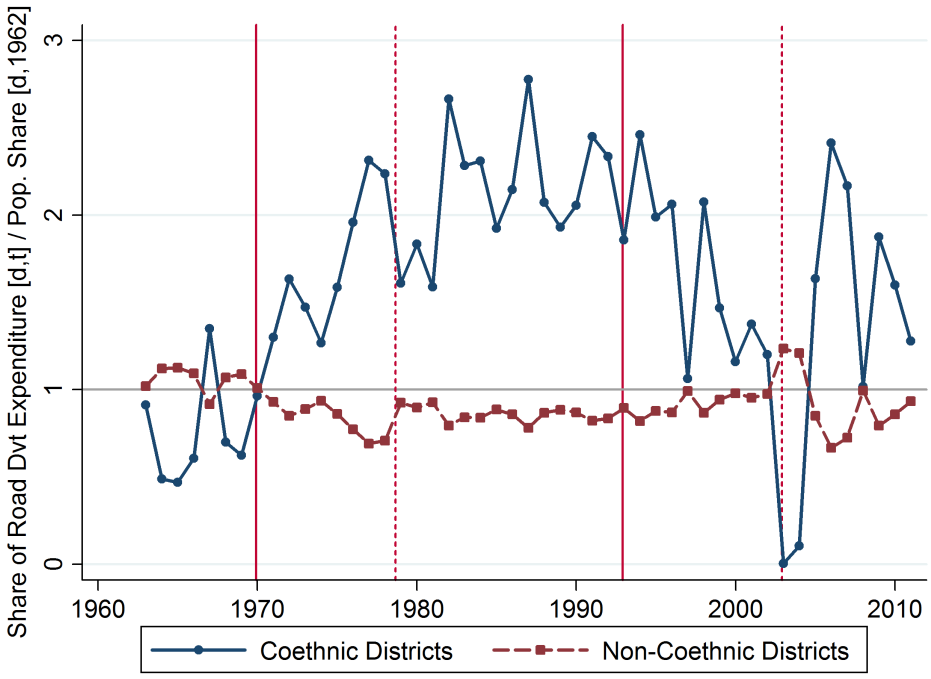


Panel B: Counterfactual Network based on Population and Distance (Market Potential)



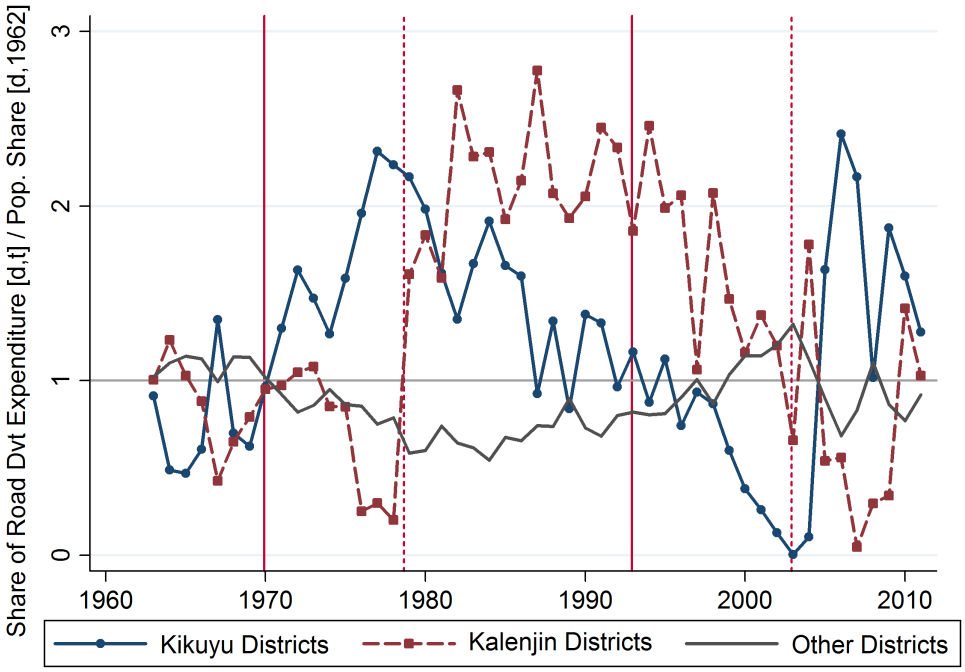
Notes: These figures illustrate the evolution of Kenya's actual and counterfactual paved road networks for key political and leadership transition years, 1969 (transition from democracy to autocracy), 1979 (from Kenyatta [Kikuyu] to Moi [Kalenjin] in 1978), 1992 (return to democracy) and 2002 (from Moi [Kalenjin] to Kibaki [Kikuyu]). The counterfactual network sequentially paves the unpaved bilateral connections with the highest market potential (based on population and distance). Border towns are not illustrated due to space constraints. Road maps are overlaid on ethnic demographics to illustrate the two presidential coethnic districts, the Kikuyus and Kalenjins. These are seven districts dominated by Kikuyus and six dominated by Kalenjins. **Coethnicity [d,t]** is defined if $\geq 50\%$ of the district's (d) population is coethnic to the president in year t. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Figure 4: Road Expenditure in Presidential Coethnic and Non-Coethnic Districts, 1963-2011



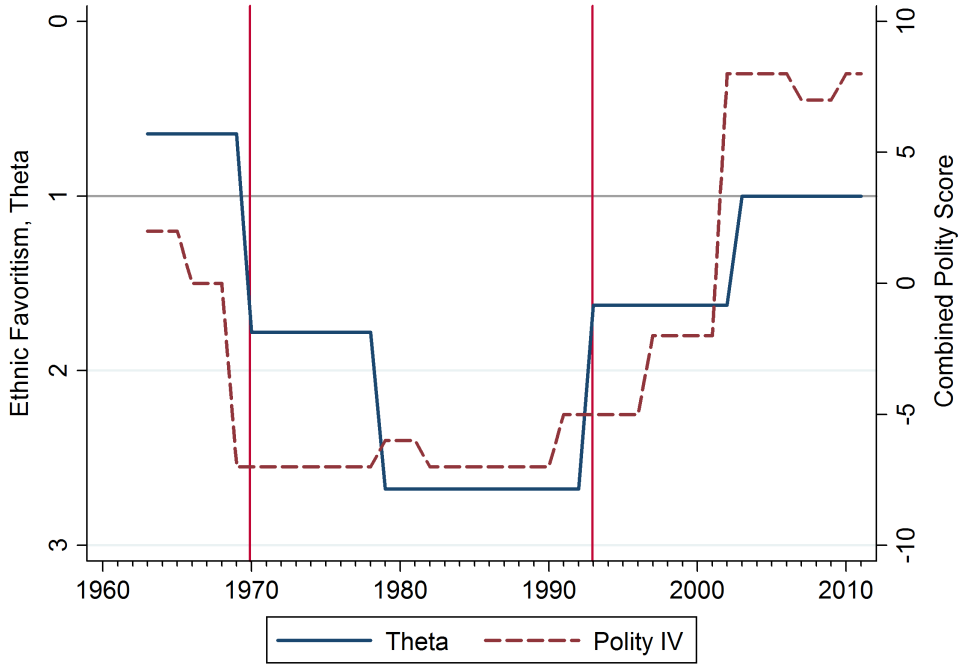
Notes: This figure plots the ratio between the share of road development expenditure in district d in year t to the share of population in 1962 for district d for coethnic and non-coethnic districts. A district d is defined as coethnic if $\geq 50\%$ of the district's population is coethnic to the president in year t. The two vertical solid red lines represent political transitions: December 1969 is the transition from democracy to autocracy, while December 1992 is the return of democracy. The two vertical red dotted lines represent leadership transitions: from Kenyatta (Kikuyu) to Moi (Kalenjin) in August 1978, and from Moi (Kalenjin) to Kibaki (Kikuyu) in December 2002. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Figure 5: Road Expenditure in Kikuyu, Kalenjin and Other Ethnic Districts, 1963-2011



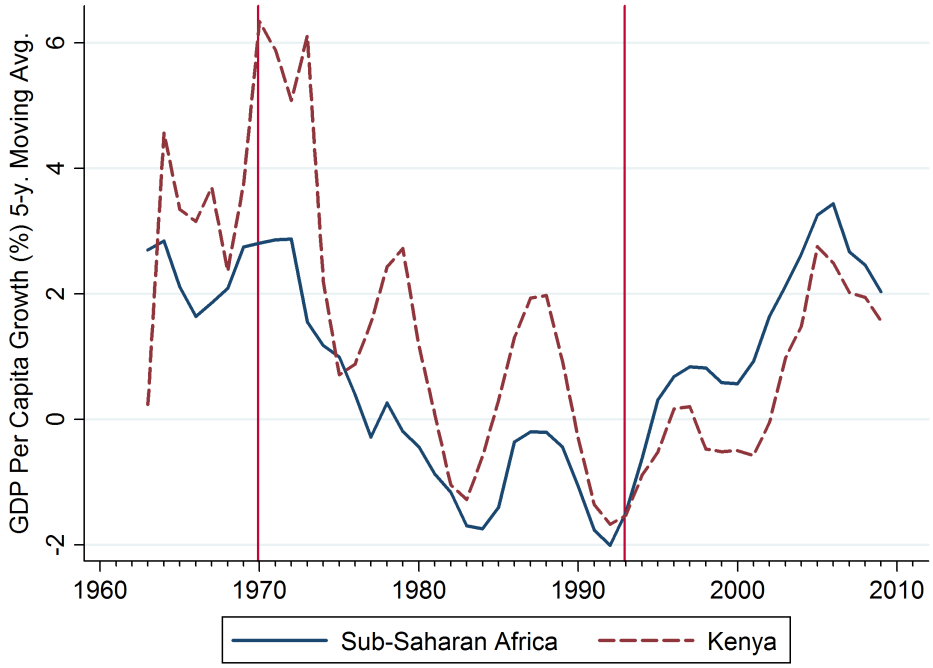
Notes: This figure plots the ratio between the share of road development expenditure in district d in year t to the share of population in 1962 for Kikuyu, Kalenjin and Other Ethnic districts. Kikuyu (and in turn Kalenjin and Other Ethnic) districts are defined as those districts if $\geq 50\%$ of the district's population is Kikuyu (and in turn Kalenjin and Other Ethnic). A Kikuyu president is in office during 1963-1978, a Kalenjin president during 1978-2002 and a Kikuyu again during 2002-2011. The vertical lines represent political transitions, while the vertical dotted lines represent leadership transitions as described in the notes of Figure 4. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Figure 6: Ethnic Favoritism and Political Regimes in Kenya, 1963-2011



Notes: This figure plots *theta*, our estimate of ethnic favoritism, and the revised combined polity score for Kenya annually from 1963 to 2011. The two vertical red solid lines represent political transitions: December 1969 is the transition from democracy to autocracy, while December 1992 is the return of democracy. Data sources and construction are described in Appendix A and Appendix E Table A2.

Figure 7: Evolution of GDP per capita growth in Sub-Saharan Africa, 1963-2011



Notes: This figure plots GDP per capita growth (%) for Kenya and the rest of Sub-Saharan Africa (population weighted average). We illustrate a 5-year moving average to reduce the year-to-year volatility in growth. The red vertical lines represent regime changes in Kenya: December 1969 is the transition from democracy to autocracy, while December 1992 is the return of democracy. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Table 1: Road Expenditure, Ethnicity and Democratic Changes in Kenya, 1963-2011

Dependent Variable	Share of road development expenditure [d,t] / Population share [d,1962]				
	(1)	(2)	(3)	(4)	(5)
Panel A					
Coethnic District [d,t]	0.97***	0.96***	0.96***	1.00***	0.97**
	(0.36)	(0.35)	(0.35)	(0.35)	(0.38)
Panel B					
Coethnic District [d,t]	1.57***	1.62***	1.64***	1.72***	1.56***
	(0.49)	(0.49)	(0.49)	(0.49)	(0.51)
Coethnic District [d,t] $*$ Democracy [t]	-1.11*	-1.24*	-1.27**	-1.32**	-1.08*
	(0.61)	(0.63)	(0.63)	(0.62)	(0.59)
F-test [*p-value*]	1.07	0.76	0.73	0.88	1.22
H_0: Coethnic + (Coethnic*Democracy) = 0	*[0.31]*	*[0.39]*	*[0.40]*	*[0.36]*	*[0.28]*
Observations	2009	2009	2009	2009	2009
Year and District fixed effects	Y	Y	Y	Y	Y
(Population, Area, Urbanization Rate)*trend	N	Y	Y	Y	N
(Earnings, Employment, Cash Crops)*trend	N	N	Y	Y	N
(Main Highway, Border, Dist. Nairobi)*trend	N	N	N	Y	N
District time trends	N	N	N	N	Y

Notes: OLS regressions on annual district-year expenditure panel dataset of 41 districts for the period of 1963-2011. **Coethnic District [d,t]** is a binary indicator equal to one if $\geq$ 50% of the district's (d) population is coethnic to the president in year t. **Democracy [t]** is a binary indicator equal to one if year t is a democratic year. Democratic years are identified as those when the constitution of Kenya allows multiple parties to contest elections. The *F-test* is for the hypothesis that coethnic and non-coethnic districts have equal outcomes under democracy. Columns (2)-(5) include initial controls interacted with a time trend (1963-2011). These controls are [i] *demographic*: district population (1962), district size (square km), and urbanization rate (1962), [ii] *economic activity*: district formal total earnings (1966), formal employment (1963) and value of cash crop exports (1965) and [iii] *economic geography*: binary indicator equal to one if district is on the Mombasa-Nairobi-Kampala corridor, binary indicator equal to one if the district borders Uganda or Tanzania and district centroid euclidean distance to Nairobi (km). Robust standard errors clustered at district level are reported in the parentheses with stars indicating *** p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Table 2: Road Building, Ethnicity and Democratic Changes in Kenya, 1964-2002

Dependent Variable	Share of paved road construction [d,t] / Population share [d,1962]				
	(1)	(2)	(3)	(4)	(5)
Panel A					
Coethnic District [d,t]	1.91**	1.94*	2.20*	3.71**	3.92*
	(0.94)	(0.99)	(1.09)	(1.69)	(2.20)
Panel B					
Coethnic District [d,t]	3.00**	3.03**	3.19**	4.26**	3.28
	(1.23)	(1.26)	(1.33)	(1.74)	(2.21)
Coethnic District [d,t] * Democracy [t]	-3.55**	-3.61**	-3.45**	-2.38*	-3.27**
	(1.38)	(1.36)	(1.32)	(1.36)	(1.39)
F-test [*p-value*]	0.44	0.49	0.10	0.98	0.00
H_0: Coethnic + (Coethnic*Democracy) = 0	*[0.51]*	*[0.49]*	*[0.75]*	*[0.33]*	*[0.99]*
Observations	410	410	410	410	410
Year and District fixed effects	Y	Y	Y	Y	Y
(Population, Area, Urbanization Rate)*trend	N	Y	Y	Y	N
(Earnings, Employment, Cash Crops)*trend	N	N	Y	Y	N
(Main Highway, Border, Dist. Nairobi)*trend	N	N	N	Y	N
District time trends	N	N	N	N	Y

Notes: OLS regressions on district-year map panel dataset of 41 districts for the period of 1964-2002. Maps are published in the following years: 1964, 1967, 1969, 1972, 1974, 1979, 1981, 1984, 1987, 1992 and 2002. We measure paved road construction as the difference between the paved length across two map years. **Coethnic District [d,t]** is a binary indicator equal to one if $\geq$ 50% of district's (d) population is coethnic to the president in year, t. **Democracy [t]** is a binary indicator equal to one if year t is a democratic year. Democratic years are identified as those when the constitution of Kenya allows multiple parties to contest elections. The *F-test* is for the hypothesis that coethnic and non-coethnic districts have equal outcomes under democracy. Columns (2)-(5) include initial controls interacted with a time trend (1964-2002). Refer to Table 1 notes for description of controls. Robust standard errors clustered at district level are reported in the parentheses with stars indicating *** p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Table 3: Counterfactual Road Building, Ethnicity and Democratic Changes in Kenya, 1964-2002

Dependent Variable	Share of paved road construction [d,t] / Population share [d,1962]		
Counterfactual Ranking	Population	Distance	Population and Distance (Market Potential)
	(1)	(2)	(3)
Panel A			
Coethnic District [d,t]	0.22	-0.56	0.67
	(0.44)	(1.21)	(1.03)
Panel B			
Coethnic District [d,t]	0.20	-0.57	0.34
	(0.52)	(1.14)	(1.08)
Coethnic District [d,t] * Democracy [t]	0.08	0.05	1.38
	(1.38)	(1.34)	(2.24)
F-test [*p-value*]	0.05	0.08	0.64
H_0: Coethnic + (Coethnic*Democracy) = 0	*[0.82]*	*[0.78]*	*[0.43]*
Observations	410	410	410
Year and District fixed effects	Y	Y	Y
Controls*trend	Y	Y	Y

Notes: OLS regressions using counterfactual dataset based on maps for the period of 1964-2002. Maps are published in the following years: 1964, 1967, 1969, 1972, 1974, 1979, 1981, 1984, 1987, 1992 and 2002. The counterfactual dataset sequentially paves the network starting with the unpaved bilateral connections of towns/cities i and j with the highest value of the ranking criteria. We use 42 towns/cities in Kenya and 7 border towns/cities in neighboring countries classified in 1962 with populations of $\geq$ 2000 habitants. This results in 1155 (=42*41/2 + 42*7) bilateral connections across these towns/cities. In column (1), the counterfactual is based on maximizing population (P_{ij}) between two bilateral pairs (P_i+P_j) to obtain the ranking of the connections. In column (2), the counterfactual is based on minimizing distance between two bilateral pairs (D_{ij}) to obtain the ranking of the connections. In column (3), the counterfactual is based on maximizing market potential, incorporating both the population and distance between two town/city pairs, $(P_i + P_j)/D_{ij}$. **Coethnic District [d,t]** is a binary indicator equal to one if $\geq$ 50% of district's (d) population is coethnic to the president in year t. **Democracy [t]** is a binary indicator equal to one if year t is a democratic year. Democratic years are identified as those when the constitution of Kenya allows multiple parties to contest elections. The *F-test* is for the hypothesis that coethnic and non-coethnic districts have equal outcomes under democracy. Columns (1)-(3) include initial controls interacted with a time trend (1964-2002). Refer to Table 1 notes for description of controls. Robust standard errors clustered at district level are reported in the parentheses with stars indicating *** p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Table 4: Road Expenditure, Ethnicity and Democratic Changes in Kenya: Political and Leadership Transitions, 1963-2011

Dependent Variable	Share of road development expenditure [d,t] / Population share [d,1962]				
Leader:	KENYATTA		MOI		KIBAKI
Regime:	Democracy 1963-1969	Autocracy 1970-1978	Autocracy 1979-1992	Democracy 1993-2002	Democracy 2003-2011
	(1)	(2)	(3)	(4)	(5)
Kikuyu District [d,1962]	-0.44	0.96**	0.66	-0.88	0.00
	(0.39)	(0.39)	(0.49)	(0.57)	(0.63)
Kalenjin District [d,1962]	-0.57	-0.17	1.88***	0.70	-0.60
	(0.41)	(0.32)	(0.66)	(1.11)	(0.57)
F-test [*p-value*]	0.15	6.92**	3.13*	2.26	0.99
H_0: Kikuyu District = Kalenjin District	*[0.70]*	*[0.01]*	*[0.08]*	*[0.14]*	*[0.33]*
Observations	287	369	574	410	369
Year fixed effects	Y	Y	Y	Y	Y
District fixed effects	-	-	-	-	-

Notes: OLS regressions on annual district-year expenditure panel dataset of 41 districts for the period of 1963-2011. Sample size varies in the column depending on duration of regime. **Kikuyu (Kalenjin) District [d,1962]** is a binary indicator equal to one if $\geq$ 50% of district's (d) population is Kikuyu (Kalenjin) according to the 1962 population census. Other ethnic groups are the comparison group. Specifications do *not* include district fixed effects (the binary ethnic indicator variables are time invariant) and set of initial controls (due to limited sample size in each column). The *F-test* is for the hypothesis that Kikuyu and Kalenjin districts have equal outcomes. Robust standard errors clustered at district level are reported in the parentheses with stars indicating *** p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Table 5: Role of the Vice-President, Cabinet Composition and Coalition Politics, 1963-2011

Dependent Variable	Share of road dvt. expenditure [d,t] / Pop. share [d,1962]		Ethnic share of cabinet [e,t] / Pop. share [e,1962]		Share of road dvt. expenditure [d,t] / Pop. share [d,1962]	
	Table 1, column 4					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Coethnic District [d,t]	1.00***	1.39**	0.65***	0.92***	1.39**	1.39**
	(0.35)	(0.58)	(0.17)	(0.08)	(0.58)	(0.58)
VP-Coethnic (District [d,t] *or* Group [e,t])		0.54		0.45***	0.54	0.54
		(0.50)		(0.13)	(0.50)	(0.50)
Panel B						
Coethnic (District [d,t] *or* Group [e,t])	1.72***	2.60***	0.64***	1.08***	1.71***	1.72***
	(0.49)	(0.71)	(0.13)	(0.31)	(0.49)	(0.49)
Coethnic (District [d,t] *or* Group [e,t]) * Democracy [t]	-1.32**	-1.63**	0.02	0.03	-1.30**	-1.32**
	(0.62)	(0.68)	(0.28)	(0.29)	(0.64)	(0.59)
VP-Coethnic (District [d,t] *or* Group [e,t])		1.46**		0.94**		
		(0.61)		(0.34)		
VP-Coethnic (District [d,t] *or* Group [e,t]) * Democracy [t]		-1.44**		-0.64		
		(0.61)		(0.38)		
Kamba-Luhya-Luo District [d,1962] * Democracy [t]					0.20	
					(0.44)	
Non-Coethnic Majority < 80% [d,1962] * Democracy [t]						0.02
						(0.67)
F-test [*p-value*]	0.90	2.49	5.87**	216.4***	0.89	1.00
H_0: Coethnic + (Coethnic*Democracy) = 0	*[0.35]*	*[0.12]*	*[0.03]*	*[0.00]*	*[0.35]*	*[0.32]*
F-test [*p-value*]		0.00		2.73		
H_0: VP-Coethnic + (VP-Coethnic*Democracy) = 0		*[0.98]*		*[0.12]*		
Observations	2009	2009	169	169	2009	2009
Year and (District *or* Group) fixed effects	Y	Y	Y	Y	Y	Y
Controls	Y	Y	N	N	Y	Y

Notes: Columns (1) and (2) are OLS regressions on annual district-year expenditure panel dataset (41 districts between 1963-2011). Columns (3) and (4) are OLS regressions on cabinet-ethnic year post-general elections (between 1963-2011), for the 13 main ethnic groups [e,t]. **Coethnic District [d,t]** and **Democracy [t]** are as defined in previous tables. **Coethnic Group [e,t]** is a binary indicator equal to one if the president in year t belongs to ethnic group e. **VP-Coethnic Group [e,t]** is a binary indicator equal to one if the vice-president in year t belongs to ethnic group e. **VP-Coethnic District [d,t]** is a binary indicator equal to one if $\geq$ 50% of the district's (d) population is from the ethnic group of the vice-president in year t. **Kamba-Luhya-Luo District [d,1962]** is a binary indicator equal to one if $\geq$50% of the population of district (d) is either from the ethnic group of Kambas (2 districts), Luhyas (3 districts) or Luos (3 districts) according to the 1962 population census. **Non-Coethnic Majority <80% [d,1962]** is a binary indicator equal to one if the main ethnic group in district (d) accounts for <80% of its total population, excluding the Presidential coethnic groups. The *F-test*'s are for the null hypothesis that Presidential/Vice-Presidential coethnic districts (or groups) and Presidential/Vice-Presidential non-coethnic districts (or groups) have equal outcomes during democracy. Columns (1)-(2) and (5)-(6) include the same set of controls as in Table 1 interacted with a time trend. Columns (3) and (4) include ethnic group time trends. Robust standard errors clustered at district (or ethnic group) level are reported in the parentheses with stars indicating ***p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Table 6: Economic Growth, Ethnic Diversity and Democratic Changes Across Countries, 1960-2010

Dependent Variable	Growth of Per Capita Real GDP [c,t]			
Sample:	World	World	World	Africa
Decades:	1960s-1980s	1960s-2000s	1960s-2000s	1960s-2000s
	(1)	(2)	(3)	(4)
Ethnic [c,1960]	-0.017***	-0.006	-0.015*	-0.023*
	(0.006)	(0.005)	(0.008)	(0.011)
Ethnic [c,1960] * Democracy [c,t]			0.013	0.036**
			(0.009)	(0.015)
Democracy [c,t]			-0.002	-0.018
			(0.005)	(0.012)
F-test [*p-value*]			0.10	0.05
H_0: Ethnic + (Ethnic*Democracy) = 0			[0.76]	[0.83]
Observations	312	528	500	182
Controls	Y	Y	Y	Y

Notes: OLS regressions on country-decadal panel dataset of 110 countries (c) for the last five decades (t): 1960s, 1970s, 1980s, 1990s and 2000s. **Growth of Per Capita Real GDP [c,t]** is the average growth rate of real per capita GDP for country c in decade t. We use Easterly and Levine's (1997, henceforth EL97) dataset for the 1960s, 1970s and 1980s and extend the data for two additional decades using *Penn World Tables v7.1*. **Ethnic [c,1960]** is the index of ethnolinguistic fractionalization of country c in 1960 (EL97). **Democracy [c,t]** is a binary indicator equal to one if country c is not an autocracy in decade t, with autocracy being defined as the average combined polity score for decade t if the threshold defined by Polity IV is $<$ -5. EL97 uses seemingly unrelated regressions; we run OLS regressions with clustered standard errors at the country level, results are robust to the different method. Columns (1)-(3) include identical controls as in Table IV, column (1) of EL97 ("Indicator for the 1960s", "Indicator for the 1970s", "Indicator for the 1980s", "Indicator Variable for Sub-Saharan Africa", "Indicator Variable for Latin America and the Caribbean", "Log of Initial Income", "(Log of Initial Income) Squared", with the exception of "Log of Schooling", due to a large number of missing observations for Africa. In column (4), we restrict the sample to only the African countries (38) and run the same specification as in column (3) excluding region fixed effects. The *F-test* is for the hypothesis that more or less ethnically diverse countries have equal outcomes under democracy. Robust standard errors clustered at country level are reported in the parentheses with stars indicating ***p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

NOT FOR PUBLICATION
Web Appendices

The Value of Democracy: Evidence from Road Building in Kenya

***by* Burgess, Jedwab, Miguel, Morjaria & Padró i Miquel**

July 2014

Appendix A: Data Sources

Appendix B: Constructing Counterfactual Road Networks

Appendix C: Theory

Appendix D: Additional Figures

Appendix E: Additional Tables

NOT FOR PUBLICATION

APPENDIX A: DATA SOURCES

This appendix describes all of the data sources used in the paper. Summary statistics are reported in Appendix E: Table A2.

1. Road Expenditure Data

We construct an annual district-year road development expenditure (in constant 2000 US$) panel data set for the period 1963-2011. The *Development Estimates* of Kenya an official government publication, allows us to track road development expenditure at the district level.[1] The *Development Estimate*'s document programmatically road projects (for instance, project *Thika Main Road* is from *Thika* town to *Nyeri* town via *Limuru* town) and their actual cost.[2] When a road project spans more than one district, we use GIS tools to lay out the road segment in question and calculate the length of kilometers within each district. Hence, for projects which span multiple districts the expenditure share is weighted by distance. For the period 1963-1973, road development estimates do not document individual road projects, instead, only large nation-wide road programs are reported. We supplement our data with the government's four-year development plans and the World Bank's road project documentation to construct comparable, project level expenditure data for the 1963-1973 period.[3]

2. Road Construction Data

We create a district-map year paved road construction panel data set by constructing a GIS database of the Kenyan road network for the years we have road maps. To construct the GIS road network, we first use as a baseline the most recent GIS database that contains contemporary roads (*Global GIS*). We then proceed with our series of historical road maps to recreate the evolution of the road network in GIS.[4] Our maps limit

[1] Government of Kenya (a, 1963/64-2010/11): road expenditure is reported in East African pounds (1963-1966), Kenyan pounds (1967-1999), and Kenyan shillings (2000-present). We use Officer (2009) and IMF (2011), to convert these amounts to current US$ and use a US$ deflator series to convert to constant 2000 US$.

[2] We supplement our primary source of Government of Kenya (a, 1963/64-2010/11) with additional ministerial reports, Government of Kenya (2007/2008–2011/2012) to ensure we have accounted for all projects. Medium Term Expenditure Framework Reports available from `www.treasury.go.ke`, accessed on December 2012.

[3] We use the Government of Kenya's four year development plans for the following years: 1964-1966, 1966-1970, 1970-1974 and 1974-1978. Road construction programs in Africa during that period were primarily fully or partially funded under International Development Agency (IDA) financing program of the World Bank. We collate all the Road Program Operational Reports (available on `http://www.worldbank.org/projects`, accessed on November 2011) and these assist us to provide the sub-projects and their relative costs for the period 1963-1973.

[4] We use the road map series published by Michelin (1964-2002) for the years 1964, 1967, 1969, 1972,

us to consistently trace only the evolution of paved roads. The evolution of non-paved roads (classified as improved, laterite and dirt roads) provides a challenge due to inconsistent categorization and definitional changes in the legends of the Michelin maps. We use categories defined in the Michelin map as motorways and hard-surfaced roads as our measure of paved roads. Using GIS tools we splice the road network for the respective years with the 1964 district boundaries to create a paved road length (in kilometers) district-year panel data set of 451 observations (41 districts tracked for 11 years that we have maps). While these maps provide the stock of roads built by that particular map year, we are interested in measuring new road construction. We obtain this by taking the difference of the road length between two subsequent maps. Hence, the number of observations in our road building analysis is 410 (=451-41).

The details on how we construct the counterfactual road network data sets are described in Appendix B. Briefly, for both the paved road construction and the road development expenditure series, we use three different methodologies to create three different counterfactual datasets. The counterfactual datasets rely on two primary datasets: (i) the constructed GIS data of the paved and non-paved road network just at the turn of independence from the Michelin Map of 1964 and (ii) the population distribution of towns/cities in Kenya (42) and in the neighboring countries (7). We identify a town/city from the population census definition of urban settlements above and equal to 2000 inhabitants. For Kenya we use the 1962 population census (Government of Kenya 1965) and for the neighboring countries the nearest census year available to 1962 and interpolate.[5]

3. Ethnic Census

We use the population census of 1962 (Government of Kenya 1965) to obtain our district ethnic demographics. We scan, digitize and geo-reference a 1963 district map which allows us to construct for each district (41) its ethnic demographics by linking district names across the map and the census. The population census reports 41 ethnic classifications. In line with studies on the politics of Kenya, we aggregate the ethnic classifications into 13 groups.[6] **Coethnic District [d,t]** is a binary indicator equal to 1 if $\geq 50\%$ of district d's population is coethnic to the president in year t. The coethnicity of the president evolves as follows: between 1963-1978 the president is Kenyatta and the ethnic group is Kikuyu, between 1979-2002 the president is Moi and the ethnic group is

1974, 1979, 1981, 1984, 1987, 1992 and 2002 (i.e. 11 maps). Two additional maps were published during our study period, in 1989 and 2010, unfortunately both of these maps are an exact re-print of the 1987 and 2002 editions, respectively.

[5]The border towns/cities in the neighboring countries are as follows: Yabelo (Ethiopia), for which we use Ethiopia's population censuses of 1956 and 1967; Afmadu (Somalia), for which we use Somalia's population censuses of 1953 and 1963; Kapoeta (Sudan, now in South Sudan) for which we use Sudan's population censuses of 1955 and 1966; Moshi, Arusha and Musoma (all in Tanzania) for which we use Tanzania's population censuses of 1957 and 1967; Tororo (Uganda) for which we use Uganda's population censuses of 1959 and 1969.

[6]Kikuyu, Kalenjin, Kamba, Luo, Luhya, Maasai, Coastal, Embu, Kisii, Meru, Somali, Turkana-Samburu and Other (which are Other Africans, Arabs, Asians, Non-Africans).

Kalenjin and between 2003-2011 the president is Kibaki and the ethnic group is Kikuyu. **Democracy [t]** is a binary indicator equal to 1 if t is a democratic year. Democratic years are identified as those when the constitution of Kenya allows multiple parties to contest elections. The variable takes the value of 1 during the period 1963-1969 and 1993-2011 (both inclusive) and 0 in the interim periods.[7] **Kikuyu District [d,1962]** (**Kalenjin District [d,1962]**, **Kamba District [d,1962]**, **Luhya District [d,1962]** and **Luo District [d,1962]**) is a binary indicator equal to one if $\geq 50\%$ of district d's population is Kikuyu (Kalenjin or Kamba or Luhya or Luo, respectively) according to the 1962 population census. The **Coethnic Group [e,t]** is a binary indicator equal to one if the president belongs to ethnic group e in year t. The **VP-Coethnic District [d,t]** is a binary indicator equal to one if $\geq$50% of district d's population is from the ethnic group of the vice-president in year t. The **VP-Coethnic Group [e,t]** is a binary indicator equal to one if the vice-president belongs to ethnic group e in year t. **Non-Coethnic Majority <80% [d,1962]** is a binary indicator equal to one if the main non-coethnic group in district d accounts for <80% of its total population. Appendix Table A1 (Panel A) provides the national population share of the major ethnic groups across post-independence Kenya. The data tabulated on ethnic composition was obtained from all published population censuses (1962, 1969, 1979, 1989 and 2009). The 1999 population census did not disclose the ethnic demographics.

4. Demographic and Socioeconomic Variables

We use various archival sources to construct three sets of control variables at the district level: (i) *demography*: district population and urbanization rates are obtained from the population census (1962) and district area is estimated using GIS tools, (ii) *economic activity*: the Statistical Abstracts of Kenya (Government of Kenya 1963-66) are used to construct total formal district employment (1963) and total formal district earnings (1966) in constant 2000 US$, the value of cash crop exports is constructed using the Government of Kenya (1964) which provides reports of cash crop production for the year 1964/65[8], (iii) *economic geography*: GIS tools are used to create a binary variable which takes the value of one if the district is on the Mombasa-Nairobi-Kampala highway corridor, another binary variable is created which takes the value of one if any part of the district borders Tanzania or Uganda, the two main trading partners. Lastly the euclidean distance between the district centroid and the national capital, Nairobi, is calculated.

[7] Note Kenya's fiscal year is from July to June, *Development Estimates* for year t provide expenditure for the period July $t-1$ to June t. Moi takes presidency from 1979 (fiscal cycle July 1978-June 1979) and Kibaki takes presidency from 2003 (fiscal cycle July 2002-June 2003). Similarly, the transition to autocracy in November 1969 is considered from 1970 (fiscal cycle July 1969-June 1970) and the transition to democracy took place in December 1992 and is considered from 1993 (fiscal cycle July 1992-June 1993).

[8] The data is reported in Kenyan Shillings, using Officer (2009) and IMF (2011), we convert these amounts to current US$ and deflate the series to obtain figures in constant 2000 US$. The 1965 export price in constant 2000 US$ (FAO 2011) is used to calculate the district's total value of cash crop exports in 1965.

6. Cabinet Composition

We source archival data and construct a panel dataset of the ethnicity and position of all cabinet members between the years 1963 and 2011, after every general election (13 cabinets). This allows us to track the evolution of each ethnic group's representation in politics. We use two primary sources of data to compile this: the official listing *The National Assembly: List of Members*, *Organization of the Government of Kenya*, and Middleton (2007).[9] While the ethnicities of prominent cabinet members is well-known, information on other politicians is obtained by consulting several secondary sources and triangulating. We use: (i) the *Weekly Review* magazine, which would often discussed the ethnicity of cabinet members after each election, (ii) research done by political scientists on Kenya, especially Hornsby (1985) and Ahluwalia (1996), and (iii) direct assistance by several journalists from the top dailies in Kenya. Combining all these sources allows us to calculate the cabinet's ethnic representation in a particular year. Appendix Table A (Panel B) tabulates the evolution of the ethnic share across the political history of Kenya.

7. Electoral Data

Electoral data for the 1992 multiparty elections are obtained from the National Election Monitoring Unit (1993). Election results are tabulated at the constituency level (188), we overlay a digital geo-referenced map of constituencies (sourced from Morjaria 2014) on the geo-referenced district map (41) to allow aggregation of election results to the district level. We focus on the presidential elections. Electoral data are tabulated for each constituency and the number of votes won by each party that stood for elections is reported. We construct two variables: (i) **Margin of Victory [d,1992]**, the difference between the voting shares (%) of the winner and the runner up parties in district d and (ii) **Party Competition Herfindhal Index [d,1992]**, the Herfindahl index of voting shares of all the parties competing in district d.

8. Newspaper Articles

For the two main daily newspapers in Kenya (*The Daily Nation* and *The Standard*) which were in circulation both before and after the arrival of democracy in 1992 we employed a team of Kenyan journalists (supervised by one of the authors) to read 25 years worth of the daily editions of these two papers (i.e. close to 18,250 newspapers when we include both titles). These archives are not digitized and are in the form of microfiche and hard copies and so *LexisNexis* searches and the like were not an option. The task set for the team of journalist was to read through and catalogue whether or not a story pertaining to roads was in each of these daily editions across the 1985-2010 period. Note the journalists were not aware of our research hypotheses.

[9] Government of Kenya (b, 1963/64-2010/11).

9. Growth, Ethnic Diversity and Democracy

Data on political regimes in Sub-Saharan Africa is obtained from the Polity IV Project. We use the variable **Combined Polity Score** which takes values from −10 (hereditary monarchy) to +10 (consolidated democracy). Polity IV categorizes regimes into *autocracies* (−10 to −6), *anocracies* (−5 to +5) and *democracies* (+6 to +10). In the analysis used in this paper we classify all regimes that are not autocracies as democracy, i.e. we add anocracies (imperfect democracies) and democracies (mature democracies). The average combined polity score for Sub-Saharan Africa is computed using the individual polity scores and weighted by the population of each country obtained from World Bank (2011). GDP per capita growth in Sub-Sahara Africa is obtained from World Bank (2011).

For Table 7, we obtain Easterly and Levine's data and append their decadel dataset with two additional decades, the 1990s and 2000s using identical sources as mentioned in Easterly and Levine (1997).[10] The variables updated are, *initial income* and *annual GDP per capita*, they are both obtained from *Penn World Tables 7.1*. Annual GDP per capita is used to calculate the growth of per capita real GDP. **Democracy [c,t]** is a binary indicator equal to one if country c is not an autocracy in decade t, specifically if the average combined polity score for the whole decade t is $\geq$-5. **Ethnic [c,1960]** is obtained from Easterly and Levine (1997) and is the ethnolingustic fractionalization of country c in 1960.

Additional References for Appendix A

Ahluwalia, P. (1996) *Post Colonialism and the Politics of Kenya*, New York: Nova Science.

FAO. (2011) FAOSTAT, Rome: Food and Agricultural Organization.

Government of Kenya. (1963-66) *Statistical Abstracts of Kenya*, Nairobi: Government Printers.

Government of Kenya. (1964) *Development Plan of Kenya*, 1964-1970, Nairobi: Government Printers.

Government of Kenya. (a, 1963/64 - 2010/11) *Development Estimates for Year 19../..*, Nairobi: Government Printers.

Government of Kenya. (b, 1963/64 - 2010/11) *The National Assembly: List of Members, Organization of the Government of Kenya*, Nairobi: Government Printers.

[10] Available on `http://williameasterly.org/academic-work`, accessed on December 2012.

Government of Kenya. (1965) *Housing and Population Census of 1962*, Nairobi: Government Printers.

Government of Kenya. (2007/08-2011/12) *Physical Infrastructure Sector: Medium Term Expenditure Framework*, Nairobi: Government Printers.

Heston, A., R. Summers, and B. Aten. (2012) *Penn World Tables Version 7.1*, Centre for International Comparisons of Production, Income and Prices, Philadelphia: University of Philadelphia.

Hornsby, C. (1985) The Member of Parliament in Kenya, 1969-1983: the Election, Background and Position of the Representative and the Implications for his Role in the One-Party State, Unpublished thesis, Oxford University.

IMF. (2011) *International Financial Statistics*, Washington DC: International Monetary Fund.

Michelin. (1964-2002) *Michelin Map Africa Central & Southern Africa: 746 Series*, Paris: Michelin.

Middleton, J. (2007) *Encyclopedia of Africa South of the Sahara*, New York: C. Scribner's Sons.

Morjaria, A. (2014) "Is Democracy Detrimental for the Environment in Developing Countries? Evidence from Kenya.", Working Paper, Harvard University.

National Election Monitoring Unit. (1993) *The Multi-Party General Elections in Kenya 29 December 1992: The Report of the National Election Monitoring Unit*, Nairobi: NEMU.

Ng'weno, H. (1979-1999) *Weekly Review*, Nairobi.

Officer, L. (2009) *Exchange Rates between the United States Dollar and Forty-one Currencies*, Measuring Worth Project.

World Bank. (2011) *World Development Indicators*, Washington DC: World Bank.

NOT FOR PUBLICATION

APPENDIX B: CONSTRUCTING COUNTERFACTUAL ROAD NETWORKS

This appendix describes the steps we undertake to construct the counterfactual panel datasets that are used in **Table 3** and **Figure 3** (paved road construction) and **Appendix Table A4** and **Appendix Figure A2** (road development expenditure). We set out the details first for paved road construction and then discuss the series for road development expenditure.

Step #1: Calculating the kilometers of paved road that were constructed. In our counterfactual exercise we take as given the total length of roads constructed in each between-maps period. Availability of maps restricts how many years of paved road data we have available, in particular recall that we have maps for the years: 1964, 1967, 1969, 1972, 1974, 1979, 1981, 1984, 1987, 1992 and 2002. Digitizing the maps allows us to compute the number of kilometers of paved roads that were built between 1964 and 2002 (total paved road constructed in length is 5286 kilometers). Appendix Table A9 tabulates the kilometers of paved road that were built in each between-maps period. For instance, between 1964-1967, 212 kilometers of roads were paved, between 1967-1969, 386 kilometers of roads were paved, etc. This exercise gives us two pieces of information (i) the total number of paved roads constructed for the period 1962-2002 and (ii) the number of kilometers that are paved between available map years. The counterfactuals take these kilometers of paved road as given (this can be thought of as our *budget constraint*) and proceeds to allocate them according to efficiency criteria that we spell out below.

Step #2: Creating the set of potential road segments. Appendix Figure A6 illustrates the distribution of the paved and unpaved road networks in Kenya at independence (1964). Since all the paved roads constructed during the period 1964-2002 were initially unpaved and were already in existence in 1964, the unpaved network in 1964 indicates all the road segments that could potentially be paved. Our counterfactual simulation exercise sequentially paves these unpaved segments depending on the ranking criteria we outline below. The key advantage of using the entire road network in 1964 is that it allows us to take account of the physical geography of the country. For an illustration, the unpaved roads unambiguously circumvent Mount Kenya, Mount Elgon, Lake Victoria and Lake Naivasha, as well as the national forest reserves.

Step #3: Data to generate efficiency criteria to allocate paved roads. We posit that a social planner would be interested in connecting pairs of towns/cities that are already economically active or have the potential to be active. Data on the aggregate incomes of the 42 towns/cities in Kenya and the 7 border towns/cities in 1964 are not available. We instead use the population of each town/city as a proxy for eco-

nomic activity.[11] We use Kenya's population census (Government of Kenya 1965) and *Google Earth* to construct a GIS database of towns/cities as defined by the census of localities with inhabitants of equal and above 2000 people.[12] To locate where these towns/cities are spatially in relation to the road network, we use our road map for 1964 (the available map closest in time to Kenya's independence in 1963) and the population censuses of the neighboring countries for towns/cities in neighboring countries (using the same 2000 inhabitants definition). We identify that there are 7 border towns/cities that we should account for in assessing which pairs of settlements should be connected. Appendix Figure A6 further illustrates the border towns/cities in the neighboring countries.

Step #4: Generating efficiency criteria using the information available in 1964. We keep using our towns/cities as proxies for local economic development. We construct various criteria that allow us to obtain values in connecting town/city pairs. The procedure is as follows, there are 42 towns/cities, which imply 861 possible town/city pairs within Kenya to connect and 7 border towns/cities thus 294 possible pairs of town/city between Kenya and its neighbors, this adds up to a total of 1155 possible pairs.[13] We rank these 1155 town/city pairs using three different criteria:

(i) The first criterion creates a measure using only the population of the two settlements and sums the two, i.e., maximizes the population sum of the town/city pair, *max* $(P_i + P_j)$, where town/city are denoted by i and j. By construction, this criterion gives precedence to segments that connect populated towns/cities.

(ii) The second criterion creates a measure using only the distance between two settlements and minimizes the sum of the two, i.e. minimizes the euclidean distance between the town/city pair, *min* (D_{ij}). By construction, this criterion prioritizes shorter roads which can be seen as a shorthand for minimizing costs.

(iii) The third criterion creates a measure using both the population and distance between two settlements, known as *market potential*. This criterion maximizes $(P_i + P_j)/D_{ij}$. By construction, those cities that are close to each other and have a large number of inhabitants have a connection with higher market potential. Appendix Table A3 displays the top 20 and bottom 20 potential bilateral connections in terms of their market potential, as well as when they become paved in the counterfactual simulation.

Step #5: Ranking all potential road segments according to the three efficiency criteria to create the three counterfactual. The three different criteria allow us to rank the 1155 potential pairs in their order of importance. The ranking order

[11] This is a common approach when limited economic activity measures exists, see for instance, De Long and Shleifer (1993) and Acemoglu, Johnson and Robinson (2002), among others.

[12] We will restrict our criteria to towns/cities in 1962 to abstract away from concerns of town/city growth due to political factors and leadership changes.

[13] The 1155 connections are obtained as follows, within Kenya 861 (42x41)/2 pairs and between Kenyan towns/cities and border towns/cities 294 (42x7) pairs.

depends on the counterfactual criteria used: population alone, distance alone, and market potential. Appendix Table A3 provides an illustration of how we go about creating the counterfactual. Upon ranking all the bilateral pairs (see *Rank* column in Appendix Table A3) we first systematically eliminate all bilateral pairs that are already connected by a paved road link in 1964, using the 1964 map as guidance. This then gives us a new ranking (see *Conditional Ranking*, in Appendix Table A3). For each ranked pair we ask the question: are the pairs already connected via the existing paved road network? If yes, we omit this connection and repeat the same procedure for the next connection in the rank. If the ranked pair is ***not*** connected, we pave this connection, using the shortest route via both the paved and unpaved road network using the 1964 map. We proceed in this manner, until we have allocated all of paved road kilometers that are available between available maps years, as discussed in **Step #1** (e.g., 212 km between 1964-1967, 386 km between 1967-1969, until we exhaust the total 5286 km between 1964-2002). In the second last column of Appendix Table A3, we illustrate when the pair gets paved in the counterfactual simulation, and the last column illustrates the number of kilometers allocated to that link.[14]

Step #6: How do we now create the counterfactual data series on paved road construction at the district level between 1964-2002? Due to the geo-spatial nature of our data, we know the length and location (and thus districts) of the road linking each town/city pair. We are able to splice these segments into kilometers of paved roads within the respective districts. This allows us to construct a data series very similar in structure to the actual paved road length data, namely, the change in the total length of paved road for each district d over time. Note that we are able to repeat **Step #5** for the other two counterfactual criteria (population alone and distance alone). This allows us to create three different counterfactual data series and hence the same dependent variable used in our main analysis (Table 2) can now we be re-computed and re-analyzed using these data. The summary statistics for these counterfactual paved road construction outcomes are reported in Appendix Table A2 (Panel D).

Step #7: How do we create the counterfactual data series for the annual development expenditure series using the information we have from paved roads constructed between 1964-2002? For the analysis on road development expenditure, we use our original road development expenditure which allow us to obtain yearly amounts of road investment. Since our counterfactual exercise is based on spatially connecting settlements and hence about reallocating paved roads between two geo-spatial points using the criteria outlined in **Step #4**, we can only construct the counterfactual series for the period 1964 to 2002 and not from 1963 to 2011, which is the time period we have available for our actual road expenditure data series. From **Step #1**, we know that we have to reallocate 5286 km of paved roads during our study

[14]For the cases when the establishment of a connection spans across two periods (for instance the 1964-1967 and 1967-1969), we allocate the segments that are closer to the largest of the two cities to the first period (e.g., 1964-1967) and the segments that are farther to the second period (e.g., 1967-1969).

period. Since we know the budget contribution for each year t to the total road development expenditure budget for the nation across the whole period 1964-2002, we can compute the km of paved roads that are to be allocated every year between 1964 and 2002, using the average cost to construct 1 km of paved road. This methodology thus assumes that the cost of constructing 1 km of paved road remains constant throughout the study period.[15] Using the expenditure series, we can thus estimate how many km of paved roads can be constructed in each year. This relies on another assumption, namely that the whole roads budget is being allocated to construction of paved roads. The World Bank Operational Reports on roads indicate that the bulk of the road budget is dedicated to constructing paved roads, but once again, the need for this reasonable but strong assumption leads us to prefer the paved road length counterfactual (described above).[16] See Appendix Table A10 for these figures.

We use our estimate of the counterfactual paved road length built in each district-map year (based on the underlying road expenditures in that year), as well as the total national km of paved roads in that year (using the same cost per 1 km constructed approximation discussed above), to compute the district's share of national road expenditures in that year. We once again apply the three ranking criteria used above (see **Step #4**) to construct three counterfactual data series on district-year road expenditures, similar to the three counterfactual series created in the paved road length counterfactual above. The summary statistics for the three counterfactual road expenditure outcomes are reported in Appendix Table A2 (Panel D).

Note: The town/city pair rankings for all six counterfactual series (three each for the road length counterfactuals and for the road expenditure counterfactuals) are available from the authors (in *MS-Excel* spreadsheet format).

Additional References for Appendix B

Acemoglu, D., S. Johnson, and J. A. Robinson. (2002) "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *The Quarterly Journal of Economics*, 117(4): 1231-1294.

De Long, J. Bradford, and A. Shleifer. (1993) "Princes and Merchants: European City Growth before the Industrial Revolution." *Journal of Law and Economics*, 36(2): 671-702.

[15] Note that the paved road length counterfactual described above does not rely on this sort of "scaling" assumption, and thus is arguably more attractive, hence our primary focus on the paved road length counterfactual in the main text.

[16] See discussion on World Bank Operational Reports in Appendix A.

NOT FOR PUBLICATION

APPENDIX C: THEORY

We start by stating more formally a few features of the model.

We assume that $R(\eta)$ is strictly increasing and concave and satisfies $R'(0) = \infty$ and $R'(\infty) = 0$.

Denote by $\omega_t \in \{A, B\}$ a state variable that captures the ethnic type of the president at time t.

Formally stated, the timing of the game, given ω_t, is as follows:

1. The president announces the policy vector $P_t = (\tau^{\omega_t}, \eta^{A\omega_t}, \eta^{B\omega_t})$
2. The citizens of group ω_t decide whether to support the leader, $s_t = 1$ or not $s_t = 0$
3. If $s_t = 1$, P_t is implemented and payoffs are realized. Next period starts with $\omega_{t+1} = \omega_t$ with probability $\bar{\gamma}$. With probability $1 - \bar{\gamma}$ the president loses power and the next president is from the other group.
4. If $s_t = 0$, the leader is immediately ousted and the transition policy vector $P = (0, 0, 0)$ is implemented. After the transition, with probability $\underline{\gamma}$ the new ruler belongs to the same group as the ousted ruler and hence $\omega_{t+1} = \omega_t$. With probability $1 - \underline{\gamma}$ the new president belongs to the other group.

We now proceed to the proof of Proposition 1.

We search for the Markov Perfect Equilibrium (MPE) of the game. Strategies can therefore only be conditioned on the payoff relevant state variables and past play within the stage game. Note that the only payoff-relevant state variable is ω_t.

Assume that $\theta < \max\{\frac{1}{\pi^A}, \frac{1}{\pi^B}\}$.

Denote by $V^i(j)$ a MPE utility for a citizen of type i starting in a subgame with a president of type j.

We proceed by backwards induction. Assume a president of type i announces $P^i = (\tau^i, \eta^{Ai}, \eta^{Bi})$.

For group i to support the policy it must be that

$$\begin{aligned} R(\eta^{ii}) - \tau^i + \bar{\gamma} V^i(i) + (1 - \bar{\gamma}) V^i(j) &\geq \underline{\gamma} V^i(i) + \left(1 - \underline{\gamma}\right) V^i(j) \\ R(\eta^{ii}) - \tau^i + \left(\bar{\gamma} - \underline{\gamma}\right)\left(V^i(i) - V^i(j)\right) &\geq 0 \end{aligned} \tag{2}$$

The President thus maximizes his instanteneous utility subject to (2) and (1).

$$\begin{gathered} \max_{\tau^i, \eta^{ii}, \eta^{ij}} \pi^i \left(\tau - \eta^{ii}\right) + \pi^j \left(\tau - \eta^{ij}\right) \\ R(\eta^{ii}) - \tau^i + \left(\bar{\gamma} - \underline{\gamma}\right)\left(V^i(i) - V^i(j)\right) \geq 0 \\ \eta^{ii} \leq \theta \left(\pi^i \eta^{ii} + \pi^j \eta^{ij}\right) \\ \eta^{ji} \geq 0 \end{gathered}$$

Note that the last constraint cannot bind: if $\eta^{ji} = 0$ then due to (1) we would have $\eta^{ii} \leq \theta\pi^i\eta^{ii}$ which directly contradicts $\theta < \max\{\frac{1}{\pi^A}, \frac{1}{\pi^A}\}$.

The first order conditions of the problem yield (λ and μ as multipliers)

$$\begin{aligned} \pi^i + \pi^j - \lambda &= 0 \\ -\pi^i + \lambda R'(\eta^{ii}) + \mu\left(\theta\pi^i - 1\right) &= 0 \\ -\pi^j + \mu\theta\pi^j &= 0 \end{aligned}$$

This solves to

$$\begin{aligned} \lambda &= 1 \\ R'(\eta^{ii}) &= \frac{1}{\theta} \\ \mu &= \frac{1}{\theta} \end{aligned}$$

which means that both constraints are binding. Since this does not depend on π^i or π^j (the only differences across groups), we have that $R'(\eta^*) \equiv R'(\eta^{ii}) = R'(\eta^{jj}) = \frac{1}{\theta}$. Also, since (2) is binding, we have

$$\begin{aligned} \eta^{ji} &= \eta^* \frac{1 - \theta\pi^i}{\theta\pi^j} \\ \eta^{ij} &= \eta^* \frac{1 - \theta\pi^j}{\theta\pi^i} \end{aligned}$$

So we can now set up the value functions

$$\begin{aligned} V^i(i) &= R(\eta^*) - \tau^i + \bar{\gamma}V^i(i) + (1 - \bar{\gamma})\, V^i(j) \\ V^i(j) &= R(\eta^{ij}) - \tau^i + \bar{\gamma}V^i(i) + (1 - \bar{\gamma})\, V^i(j) \\ V^j(j) &= R(\eta^*) - \tau^j + \bar{\gamma}V^j(j) + (1 - \bar{\gamma})\, V^j(i) \\ V^j(i) &= R(\eta^{ji}) - \tau^j + \bar{\gamma}V^j(j) + (1 - \bar{\gamma})\, V^j(i) \end{aligned}$$

and in addition we know that the two versions of (1) are binding

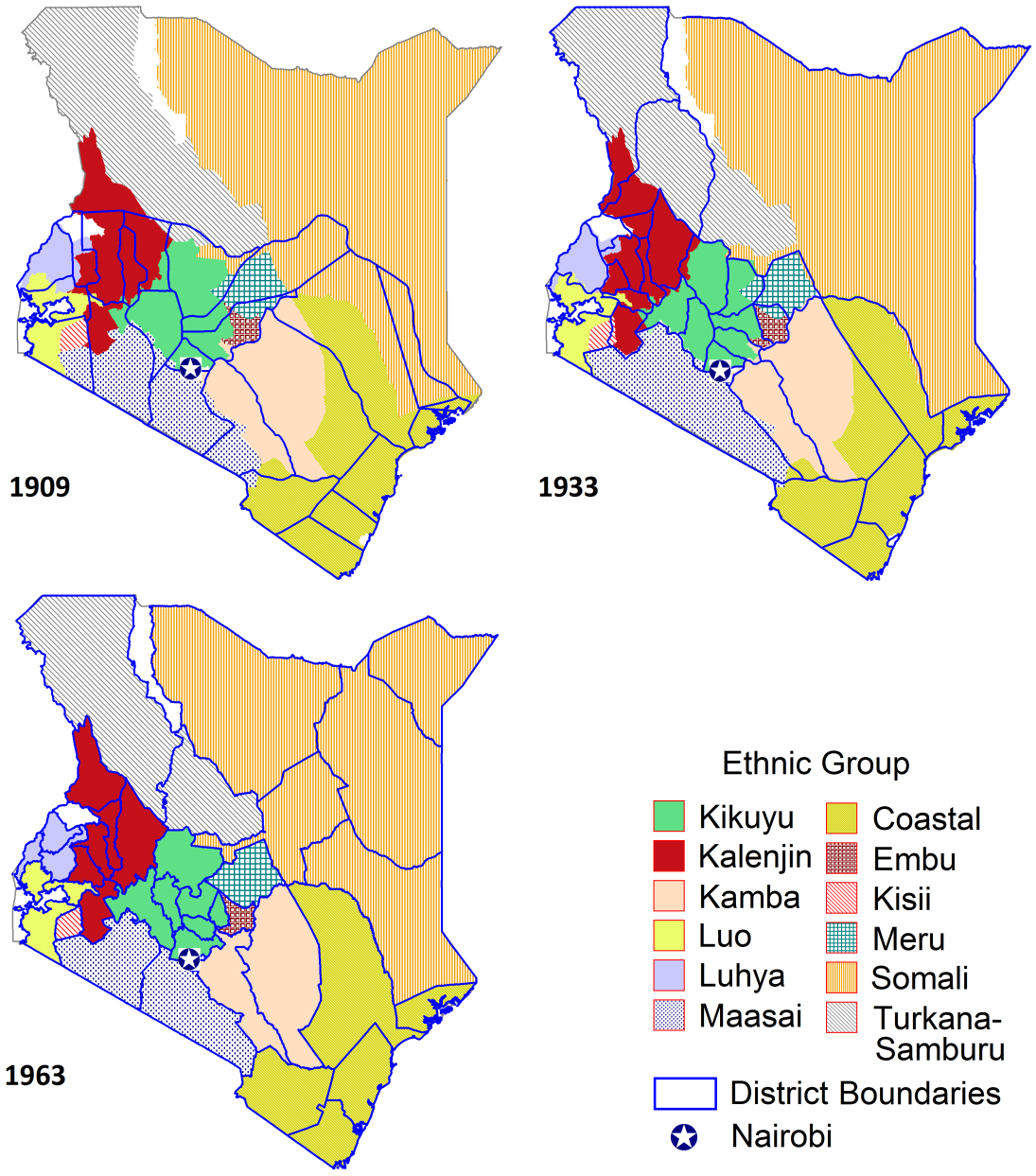
$$\begin{aligned} R(\eta^*) - \tau^i + \left(\bar{\gamma} - \underline{\gamma}\right)\left(V^i(i) - V^i(j)\right) &= 0 \\ R(\eta^*) - \tau^j + \left(\bar{\gamma} - \underline{\gamma}\right)\left(V^j(j) - V^j(i)\right) &= 0. \end{aligned}$$

This gives us a linear system of six equations in six unknowns $\left(V^i(i), V^i(j), V^j(j), V^j(i), \tau^i, \tau^j\right)$. This has a unique solution, and hence uniqueness of MPE is proven.

NOT FOR PUBLICATION

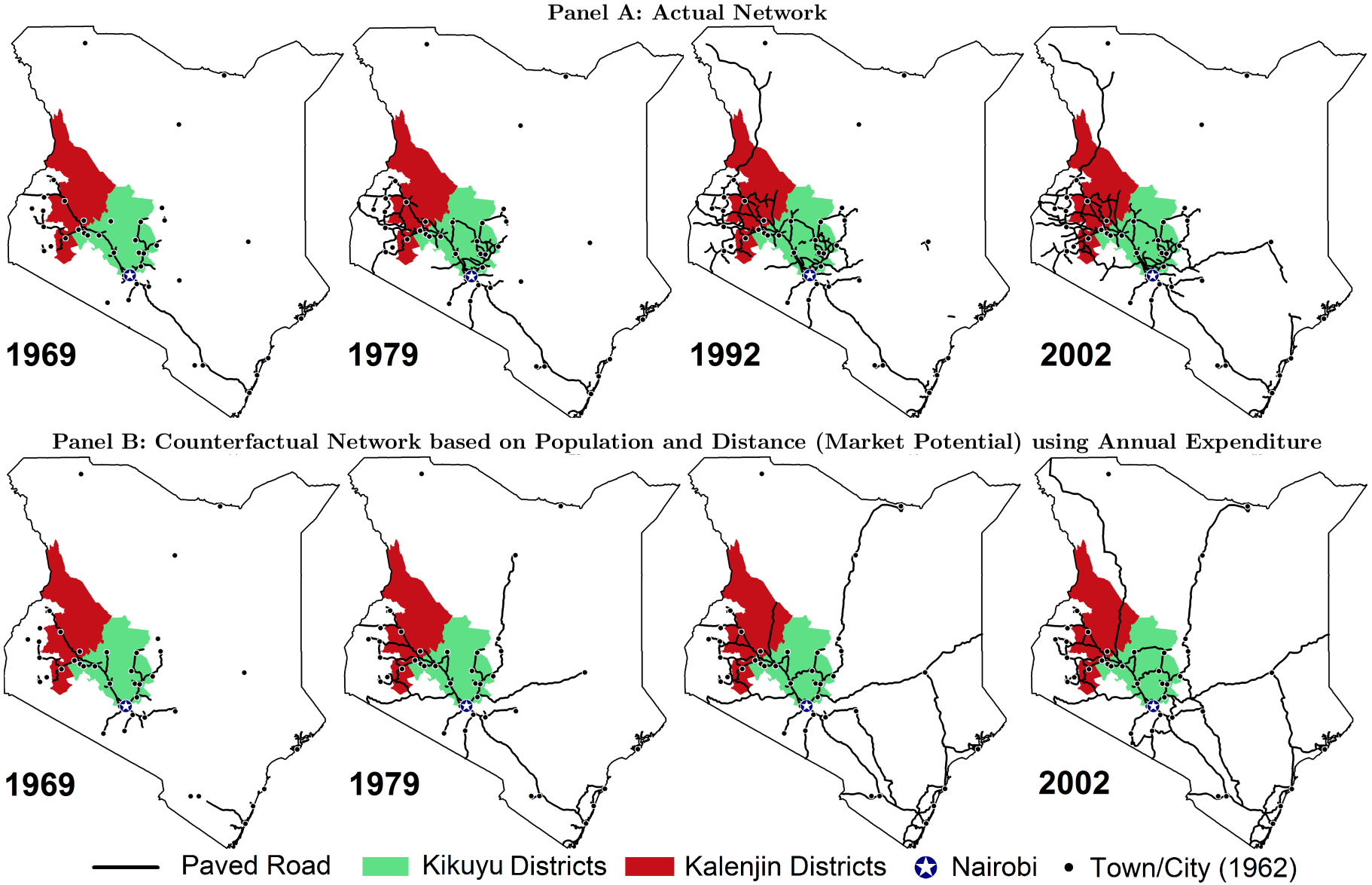
APPENDIX D: ADDITIONAL FIGURES

Appendix Figure A1: Evolving District Boundaries in Colonial Kenya and Ethnic Composition



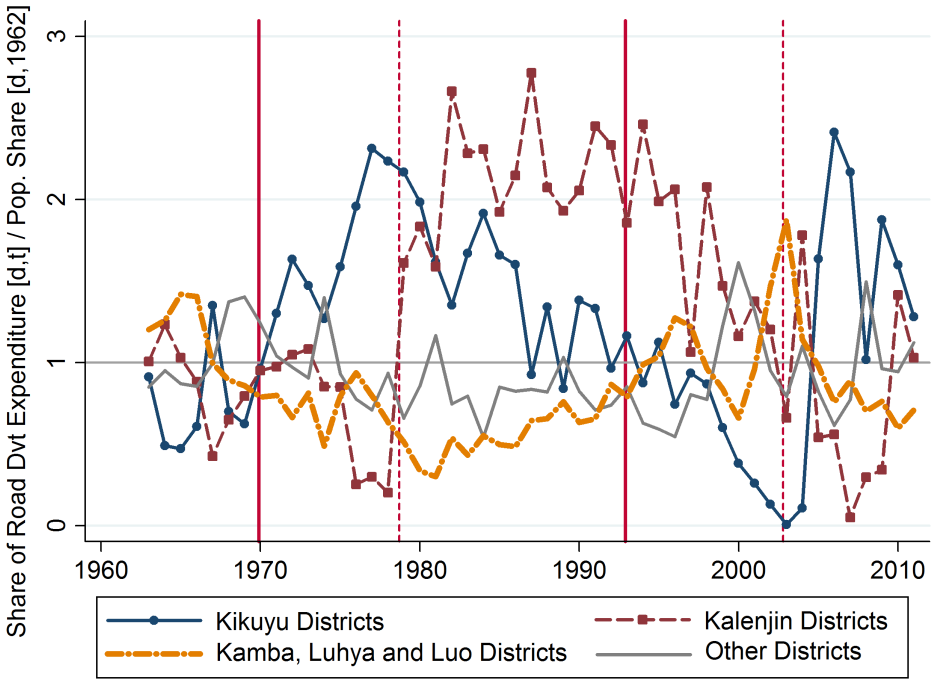
Notes: These figures illustrate the district ethnic composition, using the 1962 population census, and the evolution of district boundaries for selected years (1909, 1933, 1963) in Colonial Kenya. A district d is defined to be ethnic group e if $\geq 50\%$ of the district's population is ethnic group e. Only three districts are without a single ethnic majority group: Nairobi, Mombasa and Trans Nzoia. The 41 districts of the 1963 delineation of boundaries is used in all our analysis. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Figure A2: Evolution of Kenya's Paved Road Network



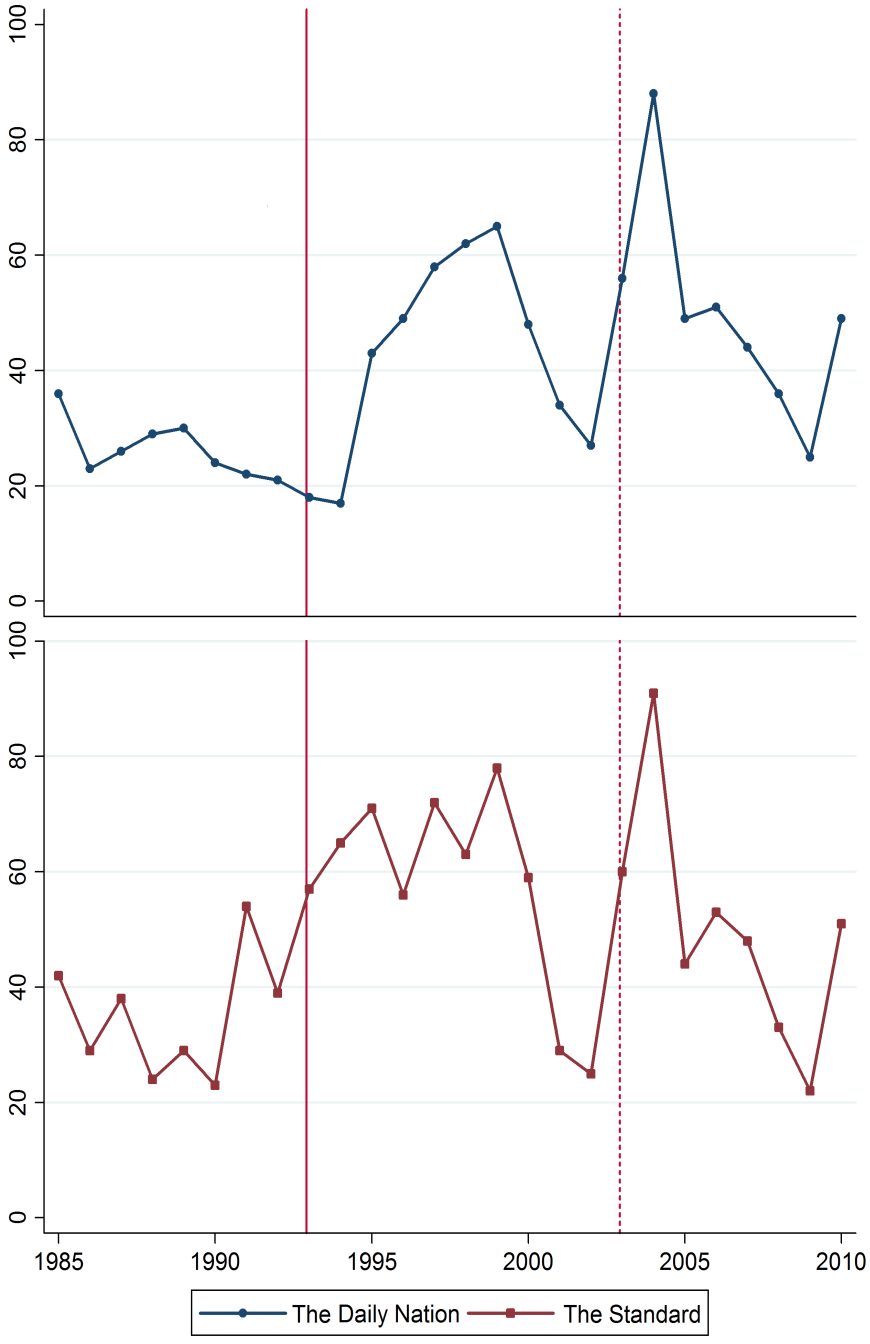
Notes: These figures illustrate the evolution of Kenya's actual and counterfactual paved road network for the key political and leadership transition years, 1969 (transition from democracy to autocracy), 1979 (from Kenyatta [Kikuyu] to Moi [Kalenjin] in 1978), 1992 (return to democracy) and 2002 (from Moi [Kalenjin] to Kibaki [Kikuyu]). The counterfactual network sequentially paves the unpaved bilateral connections with the highest market potential (based on population and distance). Border towns are not illustrated due to space constraint. The length of paved roads to be re-allocated every year is constrained by the amount of the total budget available for a year t as a share of the total road expenditure budget across the whole period of 1963-2011. Road maps are overlaid on ethnic demographics to illustrate the two presidential coethnic districts of Kikuyus and Kalenjins. **Coethnicity [d,t]** is defined if $\geq 50\%$ of district d's population is coethnic to the president in year t. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Figure A3: Road Expenditure in Kenyan Districts for the Largest Ethnic Groups and Other Groups, 1963-2011



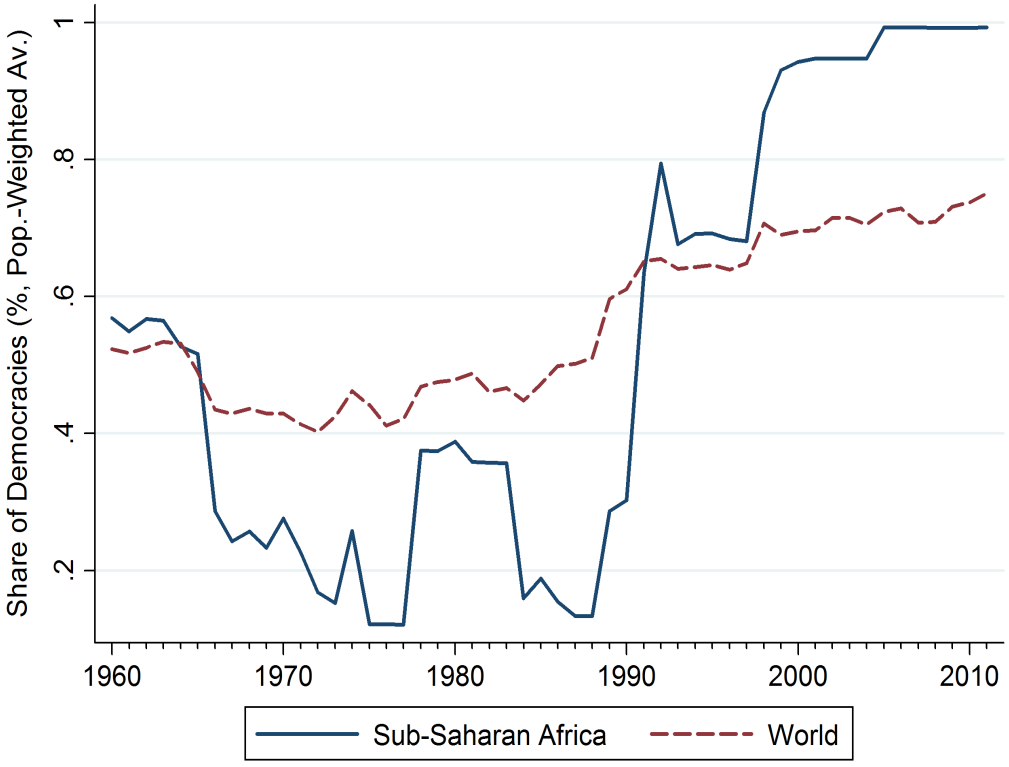
Notes: This figure plots the ratio between the share of road development expenditure in district d and year t to the share of population in 1962 for the main ethnic groups of Kenya (Kikuyu, Kalenjin and a single category for Kamba-Luhya-Luo), and the rest of the ethnic groups categorized as Other Ethnic districts. The main ethnic groups (Kikuyu, Kalenjin and Kamba-Luhya-Luo) and the Other Ethnic districts are defined as these types of district if $\geq$ 50% of the district's population is dominated by the main ethnic group or if it falls under other ethnic groups. There are 7 Kikuyu and 6 Kalenjin districts. The Kikuyu and Kalenjin districts are as defined in Figure 5. A Kamba-Luhya-Luo District is a district d if $\geq$ 50% of its population is either Kamba (2 districts), Luhya (3 districts) or Luo (3 districts) according to the 1962 population census. The vertical red lines represent political transitions, while the red vertical dotted lines represent leadership transitions as detailed in Figure 1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Figure A4: Number of Road Articles in *The Daily Nation* and *The Standard* Newspapers, 1985-2010



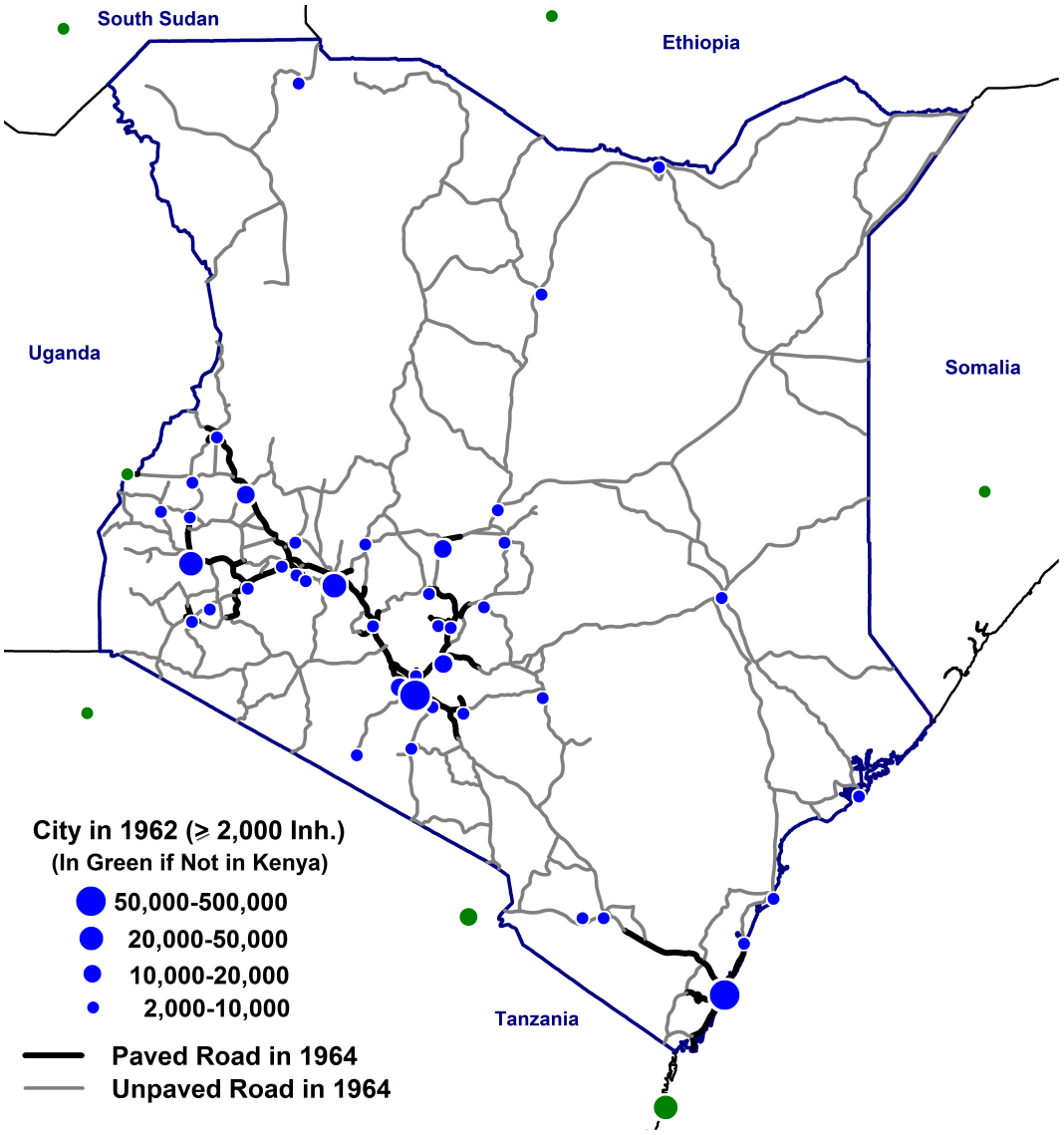
Notes: This figure illustrates the evolution of the number of articles which pertain to roads in Kenya's two leading independent dailies: *The Daily Nation* and *The Standard*. Data sources are described in Appendix A.

Appendix Figure A5: Democratic Change in the World and Sub-Saharan Africa, 1960-2011



Notes: This figure illustrates the annual share of democracies (%, pop.-weighted averages) for the world and for Sub-Saharan Africa for the period 1960-2011. A country is democratic if it is not autocratic in the Polity IV data set (combined polity score of ≥-5). Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Figure A6: Road Network and Urban Settlements in Kenya, 1962-1964



Notes: This figure illustrates the paved and unpaved road network at independence (1964) and the spatial distribution of urban settlements on the eve of independence (1962). Urban settlements are those towns/cities that have ≥2000 inhabitants. Kenya at independence has 42 towns/cities (Nairobi being the largest and the capital, followed by the port city of Mombasa). The map also depicts in green the 7 border towns/cities in neighboring countries: Ethiopia (1), Somalia (1), Sudan (1), Tanzania (3) and Uganda (1). Data sources and construction are described in Appendix A and Appendix E: Table A2.

NOT FOR PUBLICATION

APPENDIX E: ADDITIONAL TABLES

Appendix Table A1: Ethnic Group representation in Population and Cabinet, 1962-2011

***Panel A*: Population Share (%) of Main Ethnic Groups**							
Census	Kikuyu	Kalenjin	Luo	Luhya	Kamba	Other	Pop. (Millions)
1962	18.8	10.8	13.4	12.7	10.5	33.8	8.6
1969	20.1	10.9	13.9	13.3	11.0	30.8	11.0
1979	20.9	10.8	13.2	13.8	11.3	30.0	15.3
1989	20.8	11.5	12.4	14.4	11.4	29.5	21.4
2009	17.2	12.9	10.8	13.8	10.1	35.2	38.6
***Panel B*: Cabinet Share (%) of Main Ethnic Groups**							
Cabinet	Kikuyu	Kalenjin	Luo	Luhya	Kamba	Other	Cabinet Size
1963	35.3	0.0	0.0	5.9	0.0	35.3	17
1964	31.6	5.3	5.3	5.3	9.1	27.6	19
1966	27.3	4.6	4.6	9.1	9.1	36.3	22
1969	31.8	9.1	9.1	9.1	18.2	22.7	22
1974	31.8	9.1	9.1	9.1	9.1	31.8	22
1979	29.6	14.8	7.4	11.1	9.1	28.0	27
1983	20.8	16.7	12.5	12.5	0.0	37.5	24
1988	25.0	11.8	14.7	11.8	0.0	36.7	34
1993	6.0	20.0	4.0	16.0	9.1	44.9	25
1998	5.4	25.0	0.0	17.9	9.1	42.6	28
2003	21.2	7.7	15.4	19.2	18.2	18.3	26
2005	22.8	6.1	3.0	24.2	18.2	25.7	33
2008	17.4	13.9	11.6	18.6	9.1	29.4	43

Notes: Panel A tabulates the national share of the main ethnic groups for each population census. The 1999 population census did not disclose the national ethnic demographics, the national population was at 28.7 million. *Panel B* tabulates the ethnic profile of the appointed cabinet post-general elections. The cabinet includes the president, vice-president, ministers with portfolios and two other ex-officios. The solid lines in *Panel B* denote leadership transitions: from Kenyatta (Kikuyu) to Moi (Kalenjin) in August 1978, and from Moi (Kalenjin) to Kibaki (Kikuyu) in December 2002. The dashed lines in *Panel B* denote democratic regime changes in Kenya: December 1969 is the transition from democracy to autocracy, while December 1992 is the return of democracy. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Table A2: Summary Statistics and Data Sources

(Details provided in **Appendix A**)

	Mean	SD	Obs	Source
Panel A: Main Dependent Variables				
(a) Share of Road Expenditure [d,t]/Population Share [d,1962]	1.25	2.80	2009	*Expenditure*: Government of Kenya (a, 1963/64-2010/11); *Population*: Government of Kenya (1965).
(b) Share of Road Expenditure [d,t]/Area Share [d]	3.62	7.81	2009	*Expenditure*: sources in (a); *Area*: GIS tools on map of districts.
(c) Share of Paved Road Construction [d,t]/Pop. Share [d,1962]	1.37	7.29	410	*Paved Roads*: Michelin (1964-2002); *Population*: sources in (a).
(d) Share of Paved Road Construction [d,t] / Area Share [d]	2.90	8.30	410	*Paved Roads*: sources in (c); *Population*: sources in (a).
Panel B: Main Regressors				
(a) Coethnic District [d,t]	0.16	0.37	2009	*Coethnic*: Political & Leadership changes (see Figure 1); *Ethnic population*: Government of Kenya (1965).
(b) Democracy [t]	0.53	0.50	49	*Democratic years*: Political & Leadership changes (see Figure 1).
(c) Coethnic Share [d,t]	0.12	0.29	2009	*Coethnic*: Political & Leadership changes (see Figure 1); *Ethnic population*: sources in (a).
(d) Kikuyu District [d,1962]	0.17	0.38	41	*Ethnic Population*: sources in (a).
(e) Kalenjin District [d,1962]	0.15	0.36	41	*Ethnic Population*: sources in (a).
Panel C: Control Variables				
(a) Population ('000) [d,1962]	211	164	41	*Population*: Government of Kenya (1965).
(b) Area ('000 Sqr km) [d]	13.9	17.4	41	*Area*: GIS tools on map of districts.
(c) Urbanization (%) [d,1962]	7.4	20.0	41	*Ethnic Population*: sources in (a).
(d) Total Earnings in the Formal Sector (2000$) [d,1966]	10.5	21.0	41	*Earnings:* Government of Kenya (1963-66).
(e) Total Employment in the Formal Sector ('000) [d,1963]	42.6	77.2	41	*Employment*: Government of Kenya (1963-66).
(f) Total Value of Cash Crop Production (2000$) [d,1963]	8.3	20.1	41	*Cash Crop*: Government of Kenya (1964); *Value*: Officer (2009) and IMF (2011).
(g) Mombasa-Nairobi-Kampala Corridor [d]	0.37	0.48	41	*Corridor*: GIS tools on map of districts.
(h) Border district [d]	0.27	0.44	41	*Border*: GIS tools on map of districts.
(i) Euclidean Distance (km) to Nairobi [d]	268.3	146.1	41	*Distance*: GIS tools on map of districts.

Appendix Table A2: Summary Statistics and Data Sources

(continued, details provided in **Appendix A**)

	Mean	SD	Obs	Source
Panel D: Counterfactual Dependent Variables				
(a) Share of Paved Road Construction [d,t]/Population Share [d,1962]				
Counterfactual criteria based on:				
(i) population	1.86	7.95	410	*Paved Roads*: Michelin (1964-2002); *Population*: Government of Kenya (1965).
(ii) distance	1.69	8.01	410	*Paved Roads*: Michelin (1964-2002); *Location of settlements*: GIS map of towns/cities from Government of Kenya (1965).
(iii) population and distance (market potential)	2.25	11.50	410	*Paved Roads*: Michelin (1964-2002); *Population and location of settlements*: Government of Kenya (1965).
(b) Share of Road Expenditure [d,t]/Population Share [d,1962]				
Counterfactual criteria based on:				
(i) population	2.01	13.20	1599	*Paved Roads*: Michelin (1964-2002); *Population*: Government of Kenya (1965).
(ii) distance	1.76	10.67	1599	*Paved Roads*: Michelin (1964-2002); *Population and location of settlements*: Government of Kenya (1965).
Panel E: Other Main Variables				
(a) Ethnic share of cabinet [e,t]/Pop. share [e,1962]	1.03	0.80	169	*Cabinet*: Government of Kenya (b, 1963-2011), Ahluwalia (1996), Hornsby (1985), Ng'weno (1979-1999) and Middleton (2007); *Ethnic Population*: Government of Kenya (1965).
(b) Coethnic Group [e,t]	0.08	0.27	169	*Coethnic*: Political and Leadership changes (see Figure 1); *Ethnic Population*: sources in (a).
(c) VP-Coethnic Group [e,t]	0.09	0.29	169	*VP & Ethnic Population*: see sources in (a).
(d) VP-Coethnic District [d,t]	0.15	0.36	2009	*VP & Ethnic Population*: see sources in (a).
(e) Kamba-Luhya-Luo District [d,1962]	0.20	0.40	2009	*Ethnic Population*: sources in (a).
(f) Non-Coethnic Majority $< 80\%$ [d,1962]	0.37	0.48	2009	*Ethnic Population*: sources in (a).

Appendix Table A3: The Top and Bottom 20 Bilateral Connections for Construction of Paved Road Counterfactual based on Population and Distance (Market Potential)

Rank	Origin (i)	Destination (j)	Market Potential$_{ij}$ $(P_i + P_j)/D_{i,j}$	Conditional Ranking	Construction Year (in the counterfactual)	Kilometers Constructed
1	NAIROBI	KIAMBU	26310.44	$*$	–	–
2	KIKUYU	NAIROBI	22121.72	1	1967	17.5
3	NAIROBI	ATHI RIVER	13483.00	$*$	–	–
4	NAIROBI	THIKA	8810.95	$*$	–	–
5	NAIROBI	MACHAKOS	6240.04	$*$	–	–
6	NAIROBI	KAJIADO	5651.97	2	1967	51.5
7	KARURI	NAIROBI	5033.10	$*$	–	–
8	MURANGA	NAIROBI	4814.91	$*$	–	–
9	NAIVASHA	NAIROBI	4522.46	$*$	–	–
10	NAIROBI	MAGADI	3843.05	3	1967	103.1
11	KILIFI	MOMBASA	3623.13	$*$	–	–
12	NAIROBI	NYERI	3581.14	$*$	–	–
13	EMBU	NAIROBI	3188.80	4	1967	15.9
14	NAKURU	NAIROBI	2775.15	$*$	–	–
15	NAIROBI	KITUI	2643.00	5	1967 & 1969	89.1
16	NAIROBI	NANYUKI	2386.33	6	1969	50.2
17	NYAHURURU	NAIROBI	2235.42	7	1969	50.7
18	NAIROBI	ELBURGON	2207.71	8	1969	31.9
19	NAIROBI	MOLO	2067.63	$*$	–	–
20	MERU	NAIROBI	1978.64	9	1969	78.8
...	...	...	...	...	...	...
...	...	...	...	...	...	...
1136	MOYALE	NYAMIRA	6.21	*695*	NA	NA
1137	VOI	YABELO [ETHIOPIA]	6.00	*696*	NA	NA
1138	KAJIADO	LOKITAUNG	5.92	*697*	NA	NA
1139	VOI	MOYALE	5.90	*698*	NA	NA
1140	LOKITAUNG	GARISSA	5.88	*699*	NA	NA
1141	NYAMIRA	AFMADU [SOMALIA]	5.74	*700*	NA	NA
1142	LOKITAUNG	AFMADU [SOMALIA]	5.64	*701*	NA	NA
1143	GARISSA	KAPOETA [SUDAN]	5.62	*702*	NA	NA
1144	MUMIAS	KILIFI	5.50	*703*	NA	NA
1145	MUMIAS	AFMADU [SOMALIA]	5.46	*704*	NA	NA
1146	WUNDANYI	YABELO [ETHIOPIA]	5.42	*705*	NA	NA
1147	KILIFI	YABELO [ETHIOPIA]	5.25	*706*	NA	NA
1148	VOI	KAPOETA [SUDAN]	5.20	*707*	NA	NA
1149	MOYALE	WUNDANYI	5.19	*708*	NA	NA
1150	KILIFI	MOYALE	5.11	*709*	NA	NA
1151	LOKITAUNG	VOI	4.99	*710*	NA	NA
1152	WUNDANYI	KAPOETA [SUDAN]	4.75	*711*	NA	NA
1153	WUNDANYI	LOKITAUNG	4.44	*712*	NA	NA
1154	KILIFI	KAPOETA [SUDAN]	4.36	*713*	NA	NA
1155	KILIFI	LOKITAUNG	4.12	*714*	NA	NA

Notes: The above tabulation displays the top and bottom 20 bilateral connections for the construction of paved roads based on maximizing market potential (population and distance). We use data for 42 towns/cities in Kenya and 7 towns/cities in bordering countries in 1962. There are 1155 bilateral connections to consider (42*41/2 = 861 pairs within Kenya and 42*7 = 294 pairs between Kenyan towns/cities and bordering countries). Market potential for a town/city i and town/city j is defined as the sum of its population's P divided by the Euclidean distance D (km) between the pair: $(P_i + P_j)/D_{ij}$. We use the initial road network at independence (1964) to establish roads that are paved and unpaved. If a town/city pair is already connected in 1964 we ignore the pair (denoted above as $*$) and re-rank the initial ranked list, hence the conditional ranking column. If the pair has not been already paved, this pair than features in the ranking and is in line for potential paving, note this is conditional on an existence of an unpaved road connection between the town/city pair. We proceed similarly until we exhaust the amount of paved roads available (5286 km of paved roads is the total paved road constructed between 1964 and 2002, 51 town/city pairs are connected and the ranking of the last pair connected is ranked at 91 hence ranked pairs in *italics* are those that never get paved in the counterfactual as there is not enough paved roads available to allocate across the remaining years). Sudan refers to the now South Sudan. Construction of the counterfactual is described in Appendix B.

Appendix Table A4: Counterfactual Road Expenditure, Ethnicity and Democratic Changes in Kenya, 1963-2011

Dependent Variable	Share of road development expenditure [d,t] / Population share [d,1962]		
Counterfactual Ranking	Population	Distance	Population and Distance (Market Potential)
	(1)	(2)	(3)
Panel A			
Coethnic District [d,t]	0.41	0.11	0.40
	(0.70)	(0.82)	(0.49)
Panel B			
Coethnic District [d,t]	0.59	-0.16	0.11
	(0.97)	(0.95)	(0.66)
Coethnic District [d,t] * Democracy [t]	-0.52	0.79	0.85
	(1.35)	(1.31)	(1.50)
F-test [*p-value*]	0.01	0.30	0.67
H_0: Coethnic + (Coethnic*Democracy) = 0	*[0.94]*	*[0.59]*	*[0.42]*
Observations	1599	1599	1599
Year and district fixed effects	Y	Y	Y
Controls*trend	Y	Y	Y

Notes: OLS regressions on annual district-year counterfactual expenditure panel dataset of 41 districts for the period of 1964-2002. The amount of paved roads to be constructed every year is determined by the contribution of road development expenditure t to the national total road development expenditure for the whole period of 1964-2002. The counterfactual dataset paves sequentially the network starting with the unpaved bilateral connections of cities i and j with the highest value of the ranking metric. We use 42 towns/cities in Kenya and 7 border towns/cities in neighboring countries classified in 1962 with population of $\geq$ 2000 habitants. This results in 1155 (=42*41/2 + 42*7) bilateral connections across these towns/cities. In Column (1), the counterfactual is based on maximizing population (P_{ij}) between two bilateral pairs ($P_i + P_j$) to obtain the ranking of the connections. In Column (2), the counterfactual is based on minimizing distance between two bilateral pairs (D_{ij}) to obtain the ranking of the connections. In Column (3), the counterfactual is based on maximizing market potential, the population and distance between two city pairs, $(P_i + P_j)/D_{ij}$. **Coethnic District [d,t]** is a binary indicator equal to one if $\geq$ 50% of district's (d) population is coethnic to the president in year t. **Democracy [t]** is a binary indicator equal to one if year t is a democratic year. Democratic years are identified as those when the constitution of Kenya allows multiple parties to contest elections. The *F-test* is for the hypothesis that coethnic and non-coethnic districts have equal outcomes under democracy. Columns (1)-(3) include initial controls interacted with a time trend (1964-2002). Refer to Table 1 notes for description of controls. Robust standard errors clustered at district level are reported in the parentheses with stars indicating *** p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Table A5: Road Expenditure, Ethnicity and Democratic Changes in Kenya: Robustness to Excluding Selected Districts, 1963-2011

Dependent Variable	Share of road expenditure [d,t] / Population share [d,1962]				
Sample	EXCLUDING SELECTED DISTRICTS				
	White Highlands	Nairobi and Adjacent	Mombasa-Kampala	Nairobi-Kampala	Five Richest
	(1)	(2)	(3)	(4)	(5)
Panel A					
Coethnic District [d,t]	1.28***	1.08***	0.87***	0.86***	0.84**
	(0.43)	(0.36)	(0.31)	(0.30)	(0.33)
Panel B					
Coethnic District [d,t]	2.33***	1.74***	1.69**	1.97***	1.78***
	(0.79)	(0.52)	(0.63)	(0.63)	(0.59)
Coethnic District [d,t] * Democracy [t]	-2.10**	-1.23*	-1.46*	-1.98**	-1.75**
	(0.91)	(0.67)	(0.74)	(0.77)	(0.65)
F-test [*p-value*]	0.49	1.29	0.53	0.00	0.01
H_0: Coethnic + (Coethnic*Democracy) = 0	*[0.49]*	*[0.26]*	*[0.47]*	*[0.99]*	*[0.91]*
Observations	1568	1813	1274	1568	1764
No. of districts	32	37	26	32	36
Year and district fixed effects	Y	Y	Y	Y	Y
Controls*trend	Y	Y	Y	Y	Y

Notes: OLS regressions on annual district-year expenditure panel dataset of 41 districts for the period of 1963-2011. **Coethnic District [d,t]** is a binary indicator equal to one if $\geq$ 50% of district's (d) population is coethnic to the president in year t. **Democracy [t]** is a binary indicator equal to one if year t is a democratic year. Democratic years are identified as those when the constitution of Kenya allows multiple parties to contest elections. Column (1): excludes the 9 former White Highland districts (which are defined as districts which had $\geq$ 50% of it's area under the former White Highland area); column (2) excludes Nairobi and its 3 adjacent districts; Column (3) excludes 15 districts which are on the Mombasa-Nairobi-Kampala corridor; column (4) excludes the North-Western corridor connecting Nairobi to Kampala and column (5) excludes the 5 richest districts just prior to independence in 1962 (the criteria being highest African formal sector earnings at the district level). The *F-test* is for the hypothesis that coethnic and non-coethnic districts have equal outcomes under democracy. Columns (1)-(5) include the same controls as in Table 1, column 4 interacted with a time trend. Refer to Table 1 notes for description of controls. Robust standard errors clustered at district level are reported in the parentheses with stars indicating *** p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Table A6: Robustness Checks with Different Dependent Variables

Dependent Variable	Share of road dvt exp.[d,t] / Pop.sh.[d,1962]		Share of road dvt exp.[d,t] / Area sh.[d]		Share of paved road con.[d,t] / Pop.sh.[d,1962]		Share of paved road con.[d,t] / Area sh.[d]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Coethnic District [d,t]	1.00***	1.72***	1.78*	2.98***	3.71**	4.26**	4.54***	5.67***
	(0.35)	(0.49)	(0.87)	(0.99)	(1.69)	(1.74)	(1.54)	(1.68)
Coethnic District [d,t] * Democracy [t]		-1.32**		-2.22*		-2.71*		-4.80
		(0.62)		(1.29)		(1.36)		(2.89)
F-test [*p-value*]		0.88		0.46		0.98		0.11
H_0: Coethnic + (Coethnic*Democracy) = 0		*[0.36]*		*[0.50]*		*[0.33]*		*[0.75]*
Panel B								
Coethnic Share [d,t]	1.23***	2.28***	2.28**	4.36***	4.20**	5.03**	4.22**	5.76***
	(0.38)	(0.56)	(1.00)	(1.14)	(1.88)	(1.98)	(1.93)	(1.76)
Coethnic Share [d,t] * Democracy [t]		-1.88***		-3.74**		-2.85*		-5.35
		(0.66)		(1.45)		(1.66)		(4.03)
F-test [*p-value*]		0.94		0.25		1.07		0.01
H_0: Coethnic + (Coethnic*Democracy) = 0		*[0.34]*		*[0.62]*		*[0.31]*		*[0.92]*
Observations	2009	2009	2009	2009	410	410	410	410

Notes: Columns (1)-(4): OLS regressions on annual district-year expenditure panel dataset of 41 districts for the period of 1963-2011. Columns (5)-(8): OLS regressions on paved road construction panel dataset of 41 districts for the period of 1963-2002. In columns (1)-(2) and (5)-(6) the dependent variables' numerator is normalized by the population share of the district in 1962. In columns (3)-(4) and (7)-(8), the numerator is normalized by the share of the district area. All regressions include district and year fixed effects. **Coethnic District [d,t]** is a binary indicator equal to one if $\geq$ 50% of district's (d) population is coethnic to the president in year (t). **Democracy [t]** is a binary indicator equal to one if year t is a democratic year. Democratic years are identified as those when the constitution of Kenya allows multiple parties to contest elections. **Coethnic Share [d,t]** is the population share of the ethnic group of the president in district d at time t. The *F-test* is for the hypothesis that coethnic and non-coethnic districts have equal outcomes under democracy. Columns (1)-(8) include the same controls interacted with a time trend as in Table 1, column (4). Refer to Table 1 notes for description of controls. Robust standard errors clustered at district level are reported in the parentheses with stars indicating *** p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Table A7: Robustness Checks with Different Specifications

Dependent Variable	Share of road expenditure [d,t] / Population share [d,1962]				
	Table 1, column (4)				
	(1)	(2)	(3)	(4)	(5)
Panel A					
Coethnic District [d,t]	1.00***	1.25**	1.63	1.00***	1.00***
	(0.35)	(0.51)	(0.98)	(0.33)	(0.28)
Panel B					
Coethnic District [d,t]	1.72***	2.17***	1.84*	1.72***	1.72***
	(0.49)	(0.70)	(0.96)	(0.48)	(0.38)
Coethnic District [d,t] * Democracy [t]	-1.32**	-1.81	-1.26**	-1.32**	-1.32**
	(0.62)	(1.11)	(0.50)	(0.61)	(0.48)
F-test [*p-value*]	0.88	0.22	0.41	0.95	1.05
H_0: Coethnic + (Coethnic*Democracy) = 0	*[0.36]*	*[0.64]*	*[0.52]*	*[0.34]*	*[0.30]*
Observations	2009	2009	2009	2009	2009
No. of districts	41	41	41	41	41
Year and district fixed effects	Y	Y	Y	Y	Y
Controls*trend	Y	N	Y	Y	Y
Controls*year fixed effects	N	Y	N	N	N
Number of years coethnic district	N	N	Y	N	N
Clustering / Conley standard errors	District	District	District	200 km	400 km

Notes: OLS regressions on annual district-year expenditure panel dataset of 41 districts for the period of 1963-2011. **Coethnic District [d,t]** is a binary indicator equal to one if $\geq$ 50% of district's (d) population is coethnic to the president in year (t). **Democracy [t]** is a binary indicator equal to one if year t is a democratic year. Democratic years are identified as those when the constitution of Kenya allows multiple parties to contest elections. Column (1): replication of Table 1, column 4. Column (2): We interact our controls with year fixed effects. Column (3): We include the number of years a district has been a coethnic district. Column (4): Standard errors corrected for spatial clustering using a 200 km threshold. Column (5) Standard errors corrected for spatial clustering using a 400 km threshold. The *F-test* is for the hypothesis that coethnic and non-coethnic districts have equal outcomes under democracy. Columns (1)-(5) include the same controls as in Table 1, column 4 interacted with a time trend. Refer to Table 1 notes for description of controls. Robust standard errors clustered at district level are reported [for columns (1)-(3)] in the parentheses with stars indicating *** p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Table A8: Road Expenditure, Democratic Changes and Coalition Politics in Kenya, 1963-2011

Dependent Variable	Share of road expenditure [d,t] / Population share [d,1962]					
	Table 1, column (4)	ASSESSING POLITICAL COMPETITION EFFECTS				
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B						
Coethnic District [d,t]	1.72***	1.70***	1.73***	1.71***	1.69***	1.60***
	(0.49)	(0.49)	(0.49)	(0.49)	(0.47)	(0.46)
Coethnic District [d,t] * Democracy [t]	-1.32**	-1.29**	-1.34**	-1.30**	-1.27**	-1.14*
	(0.62)	(0.63)	(0.63)	(0.64)	(0.60)	(0.58)
Kamba District [d,1962] * Democracy [t]		0.68				
		(0.57)				
Luhya District [d,1962] * Democracy [t]			-0.25			
			(0.53)			
Luo District [d,1962] * Democracy [t]				0.22		
				(0.37)		
Margin of Victory [d,1992] * Democracy [t]					-0.26	
					(0.67)	
Party Competition Herfindhal Index [d,1992] * Democracy [t]						-1.17
						(1.02)
F-test [*p-value*]	0.88	0.94	0.80	0.89	0.94	1.16
H_0: Coethnic + (Coethnic*Democracy) = 0	*[0.36]*	*[0.34]*	*[0.38]*	*[0.35]*	*[0.34]*	*[0.28]*
Observations	2009	2009	2009	2009	2009	2009
No. of districts	41	41	41	41	41	41
Year and district fixed effects	Y	Y	Y	Y	Y	Y
Controls*trend	Y	Y	Y	Y	Y	Y

Notes: OLS regressions on annual district-year expenditure panel dataset of 41 districts for the period of 1963-2011. **Coethnic District [d,t]** is a binary indicator equal to one if $\geq$ 50% of district's (d) population is coethnic to the president in year (t). **Democracy [t]** is a binary indicator equal to one if year t is a democratic year. Democratic years are identified as those when the constitution of Kenya allows multiple parties to contest elections. **Kamba (*or* Luhya *or* Luo) District [d,1962]** is a binary indicator equal to one if $\geq$ 50% of district's (d) population is Kikuyu (Kamba *or* Luhya *or* Luo) according to the 1962 population census. There are 2 Kamba, 3 Luhya and 3 Luo districts. **Margin of Victory [d,1992]** is the difference between the voting shares (%) of the winner and the runner up parties at the constituency level for the presidential elections of 1992 aggregated to the district level. **Party Competition Herfindhal Index [d,1992]** is the Herfindahl index of the voting shares of all parties at the constituency level aggregated to the district level from the presidential elections of 1992. Columns (1)-(6) include the same controls as in Table 1, column 4 interacted with a time trend. Refer to Table 1 notes for description of controls. Robust standard errors clustered at district level are reported in the parentheses with stars indicating *** p <0.01, ** p <0.05, * p <0.1. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Table A9: Kilometers of Paved Roads Constructed between 1964-2002

Map Year	Next Map Year	Length (km)
1964	1967	212
1967	1969	386
1969	1972	590
1972	1974	504
1974	1979	896
1979	1981	151
1981	1984	1,149
1984	1987	220
1987	1992	209
1992	2002	969
1964	2002	5286

Notes: The table shows kilometers of paved roads constructed between every map year for the period 1964-2002. Data sources and construction are described in Appendix A and Appendix E: Table A2.

Appendix Table A10: Kilometers of Paved Roads Constructed between 1964-2002 (Road Expenditure Counterfactuals)

Year	Length (km)
1964	43
1965	42
1966	91
1967	66
1968	113
1969	134
1970	146
1971	221
1972	234
1973	236
1974	258
1975	192
1976	120
1977	146
1978	114
1979	159
1980	161
1981	145
1982	204
1983	199
1984	203
1985	159
1986	137
1987	135
1988	108
1989	164
1990	149
1991	124
1992	106
1993	60
1994	94
1995	99
1996	134
1997	115
1998	118
1999	75
2000	102
2001	69
2002	115
1964-2002	*5286*

Notes: This table shows how many kilometers of paved roads must be constructed every year for the period 1964-2002 for use in the counterfactual road expenditure exercise. In total, 5286 km of paved roads must be constructed between 1964-2002. Since we know the contribution of each year t to the total amount of road development expenditure in 1964-2002, we can back out the amount of paved roads (km) that needs to be constructed every year between 1964 and 2002 assuming constant cost per km of paved road constructed. Data sources and construction are described in Appendix A and Appendix E: Table A2.