ELSEVIER
www.elsevier.com/locate/worlddev

CrossMark

World Development Vol. 78, pp. 308–323, 2016
0305-750X/© 2015 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

http://dx.doi.org/10.1016/j.worlddev.2015.10.018

Ethnic Heterogeneity and Public Goods Provision in Zambia: Evidence of a Subnational "Diversity Dividend"

RACHEL M. GISSELQUIST [a], STEFAN LEIDERER [b] and MIGUEL NIÑO-ZARAZÚA [a,*]
[a] *UNU-WIDER, Helsinki, Finland*
[b] *German Institute for Development Evaluation/Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), Bonn, Germany*

Summary. — The "diversity debit" hypothesis – that ethnic diversity has a negative impact on social, economic, and political outcomes – has been widely accepted in the literature. Indeed, with respect to public goods provision – the focus of this article – the conventional wisdom holds that a negative relationship between ethnic heterogeneity and public goods provision is so well-established empirically that future research should abandon examination of *whether* such a relationship exists and focus instead on *why* it exists, that is, on the mechanisms underlying a negative relationship. This article challenges the conventional wisdom on empirical grounds. It demonstrates at the sub-national level strong evidence for a "diversity *dividend*" – that is, a positive relationship between ethnic heterogeneity and some measures of public goods provision, in particular welfare outcomes related to publicly provided goods and services. Building on the literature, the article draws on new analysis at district level for Zambia, using a new dataset compiled by the authors from administrative, budget, and survey data, which cover a broader range of public goods outcomes than previous work, including information on both budgetary and welfare outcomes. The article explores why relationships may differ for sub-national budgetary and welfare outcomes, considering separate models for each. Analysis shows results to be robust across a variety of alternative specifications and models. Given the more nuanced relationship between ethnic diversity and public goods provision documented, the article argues that the key task for future work is not to address why the relationship is negative, but to study under what conditions such direction holds true, and the mechanisms that underlie a diversity *dividend.* It concludes by considering key explanatory hypotheses against the Zambian data to identify promising areas for such theory development. More broadly, while the diversity debit hypothesis highlights the costs of diversity and could be interpreted as providing support for polices that minimize it, the findings in this article are consistent with a view that diversity can be good for communities, not only for normative reasons, but also because, under some conditions, it can support concrete welfare gains.
© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Key words — ethnic diversity, ethnic divisions, government performance, education, health, Zambia

1. INTRODUCTION

The provision of public goods, a key component of government performance, varies substantially across communities. It varies in terms of which goods and services are provided, how they are provided, how well, and in what amounts. This in turn can have broad implications, directly and indirectly, for economic development. A variety of structural, institutional, and cultural factors, as well as individual agency, may contribute to this variation (see, e.g., Gormley, 2007; Lijphart, 2012; Putnam, 1993). This paper focuses on one key factor emphasized in the literature on developing countries: social divisions, in particular those expressed in ethnic terms. As Banerjee, Iyer, and Somanathan (2005, p. 639) note, "the notion that social divisions undermine economic progress, not just in extremis, as in the case of a civil war, but also in more normal times" is "one of the most powerful hypotheses in political economy." High ethnic diversity as a factor that impedes economic development has received particular attention in work on sub-Saharan Africa, not only the most ethnically diverse world region but also the least developed (Ashraf & Galor, 2013a, 2013b; Easterly & Levine, 1997; Gören, 2014; Posner, 2004). [1]

The "diversity debit" hypothesis – that ethnic diversity has a negative impact on social, economic, and political outcomes – is widely accepted (Gerring, Thacker, Lu, & Huang, 2015). Indeed, with respect to public goods provision – our focus in this article – the conventional wisdom holds that a negative relationship between ethnic divisions and public goods provision is so well-established empirically that future research should abandon examination of *whether* such a relationship exists and focus instead on *why* it exists, that is, on testing hypotheses about the mechanisms underlying this negative relationship (Habyarimana, Humphreys, Posner, & Weinstein, 2007; Lieberman & McClendon, 2013).

This article makes two key contributions to the literature: First, it challenges the conventional wisdom on empirical grounds. It shows at the sub-national level strong evidence for a positive relationship between ethnic diversity and some measures of public goods provision, in particular welfare outcomes related to goods and services publicly provided. We spotlight here findings from a handful of recent studies (Gerring *et al.*, 2015; Gibson & Hoffman, 2013; Gisselquist, 2014; Singh, 2010) and draw particularly on new analysis at district level for Zambia in which we consider both government spending and a range of welfare indicators. Given that the diversity debit hypothesis is so often applied to Africa, empirical analysis on the region is highly relevant to its testing yet it has been impeded by the relatively weak data available for the region. We employ a new dataset compiled by one of the authors from administrative, budget, and survey data, which cover a broader range of public goods outcomes than previous work (see Leiderer, 2014). As a relatively stable

* We gratefully acknowledge support for this research from the United Nations University, World Institute for Development Economics Research (UNU-WIDER) and the German Development Institute/Deutsches Institut für Entwicklungspolitik (DIE). Final revision accepted: October 2, 2015.

African country in which ethnic divisions have nevertheless been salient in routine forms of politics, Zambia thus provides a valuable test case (Lindemann, 2011a, 2011b; Posner, 2005).

Second, this study contributes to theory building: Given the much more nuanced relationship between ethnic diversity and public goods provision documented in our analysis, the key task we identify for future work is not to address why the relationship is negative, but under what conditions such direction holds true. As existing theory relates principally to the diversity debit hypothesis, there is considerable scope for theory building with reference to the mechanisms underlying diversity dividends. In this study we explore several working hypotheses drawn from the literature against the Zambian data.

The next section of this article reviews the conventional wisdom and empirical evidence with regard to the diversity debit hypothesis, and situates our analysis within a set of emerging critiques. In terms of the empirical evidence, we highlight the importance of distinguishing at a minimum between national and sub-national analysis and show that when such a distinction is made, it becomes clear that evidence at the sub-national level does not give support for a diversity debit hypothesis. Building on this discussion, we spell out what we would expect to find in the Zambian data if the diversity debit hypothesis were correct, parsing different expectations for sub-national budgetary outcomes and sub-national welfare outcomes within a fiscally centralized country like Zambia. In this context, we present evidence and possible theoretical bases for a "diversity divided." The article then turns to the Zambian case. Section 3 discusses the data, and the measures of ethnic diversity and public goods provision whereas Sections 4 and 5 present the empirical model used in our analysis, and key results, respectively. We return to the question of explanation and theory building in Section 6. A final section concludes.

2. THE DIVERSITY DEBIT HYPOTHESIS: MECHANISMS, EVIDENCE, AND CRITIQUE

Literature in the diversity debit tradition points to multiple ways in which diverse communities may have negative implications for public goods provision when compared to more homogenous communities.[2] Alesina, Baqir, and Easterly (1999), perhaps the article most well-cited as evidence of a negative relationship, builds its underlying model on two key assumptions about ethnic groups. First, ethnic groups may have different preferences over what is provided, where, and how (Chandra, 2001). Because of such preferences, community members in heterogeneous areas – when compared to those in homogeneous areas – may obtain lower utility from shared public services and thus support lower contributions to their provision. Second, members of ethnic groups may have similar preferences but be prejudiced against other groups, for instance valuing public goods less because they prefer not to mix with other groups. In the Alesina *et al.* (1999) model, the average individual's utility is $u_i = g^a(1 - l_i) + c$, where g is the public good, l_i is the distance between individual i's preferred type of public good and the public good provided, and c is private consumption (see also Kimenyi, 2006).[3] The model predicts that in a majoritarian electoral system, the median distance from the median voter's ideal type (l_i^m), an indicator of the polarization of preferences, will determine the size of the public good, which will decrease as polarization increases.

A second broad argument highlighted in the literature focuses on public goods provision as a collective action problem, resolved best in situations where social capital is strong, trust levels are high, and shirkers can be punished (Olson, 1965). Because social capital, trust, and social sanctions may be weaker in ethnically diverse communities than in homogeneous ones, diverse communities are expected to be less able to resolve the collective action problems needed to provide public goods at socially optimal levels (Bahry, Kosolapov, Kozyreva, & Wilson, 2005; Khwaja, 2009; Mavridis, 2015; Miguel & Gugerty, 2005; Putnam, 2007). Relatedly, cooperation and collective action in homogeneous communities as compared to heterogeneous ones may be facilitated at a practical level by shared language and cultural norms, geographic proximity, and within-group personal connections (Deutsch, 1966; Habyarimana *et al.*, 2007).

A third line of argument highlights the role of elites, underscoring that ethnic diversity – in addition to influencing public goods provision through the preferences and actions of individual voters and community members – may also have impact through leaders and governing bodies. For instance, just as ethnic diversity may impact collective action at the level of the individual community member, it may also do so within the bodies that govern communities as members may be less able to collaborate to pass difficult legislation or more likely to deadlock in decisions over conflicting agendas. Members of government in ethnically divided polities further may owe their power to narrow ethnic constituencies rather than more broadly-based ones, which may in turn influence them to favor policies that support their own ethnic bases over others and to divert resources in economically inefficient ways (Easterly & Levine, 1997; Franck & Rainer, 2012).

Each of these arguments is consistent with a negative relationship between ethnic diversity and measures of public goods provision. The conventional wisdom holds that this negative relationship is empirically well-established, citing a number of studies (e.g., Alesina *et al.*, 1999, Alesina & La Ferrara, 2000; Easterly & Levine, 1997; Miguel & Gugerty, 2005). The priority then for future research is considered to be the mechanisms underlying this negative relationship – that is, their further elaboration and testing (Habyarimana *et al.*, 2007; Lieberman & McClendon, 2013).

(a) *Reassessing the empirical evidence*

In reviewing the empirical evidence for the diversity debit hypothesis, one of the first points to note is that little distinction has been made in the literature between empirical analyses at the national *versus* sub-national levels. In other words, the hypothesis is routinely applied in the same way to explain variation in public goods provision both across and within countries (villages, municipalities, cities), and empirical analyses at the national level are routinely cited as evidence that the hypothesis should apply at sub-national levels and vice versa. The fact that levels of government figure little in this discussion is worthy of further consideration because there are multiple reasons that we normally expect public goods provision at national and sub-national levels of government to differ.

For one, different levels of government have different roles and responsibilities, suggesting that we normally look first to the level of government with most discretion over a particular good for explanations about its provision. For instance, study of variation in the provision of education and health services at the municipal level – as a function of municipal-level factors alone – would be potentially misleading in a country where education and health policies were highly centralized at the national level, or where decisions about education were made primarily at the local level and those about health policy at the state level. In addition, in terms of budgetary outcomes, intergovernmental transfers add a level of complexity to analyses at sub-national levels. Funding for various public services may come both from national and sub-national levels, with the relative contribution of each varying not only across countries but within them as well. In the United States, for instance, federal, state, and local governments all contribute to education and the relative contribution of the federal level varies across states.

This suggests that in evaluating the empirical bases upon which the conventional wisdom rests, it is worth at a minimum separating consideration of empirical studies at national *versus* sub-national levels. Broadly speaking, empirical support for a negative relationship appears relatively robust across major studies of national-level variation, considering a range of public goods measures and specifications.[4] Easterly and Levine (1997)'s seminal analysis showed ethnic heterogeneity to be a significant factor in explaining Africa's slow growth relative to East Asia through its effect on low schooling, insufficient infrastructure, political instability, underdeveloped financial systems, distorted foreign exchange markets, and high government deficits. Posner (2004) extended this analysis, showing fractionalization among "politically relevant" ethnic groups to be associated also with internal variation in growth *across* African countries. Baldwin and Huber (2010) find support for a negative relationship in the analysis of 46 countries using an aggregate measure based on ten variables related to education, health, sanitation, infrastructure, and the regulatory framework for private sector activity; Jackson (2013) in analysis of education, drinking water, and electricity across 18 African countries; and Gerring *et al.* (2015) for human development outcomes, including child mortality, fertility, education, and wealth, across 36 developing countries.

At the sub-national level, on the other hand, the literature shows empirical support for the diversity debit hypothesis to be less clear. Gerring *et al.* (2015) in particular find support at the national level, but the opposite – a *positive* relationship between ethnic diversity and human development outcomes – at the sub-national level. Similarly, and particularly relevant for our analysis is the study of local district councils in Zambia by Gibson and Hoffman (2013) which shows ethnic fractionalization to be *positively* correlated with local government expenditures.

Sub-national analyses have also demonstrated an important distinction between *divisions* and *diversity*, providing evidence that it is ethnic division, not diversity *per se*, that is associated with a negative relationship. Miguel (2004), in particular, shows in a comparative analysis of communities in Tanzania to those in Kenya – where national identity *versus* ethnic identity is comparatively strong and vice versa – that the negative relationship holds only in the latter context, while Singh (2010) finds based on longitudinal analysis of social development in Kerala that it is the absence of a subjective sense of "we-ness," that drives negative outcomes (see also Kanbur, Rajaram, & Varshney, 2011). This is an important caveat to the diversity debit hypothesis, suggesting that a "division debit" hypothesis might be a better term.[5]

Arguably the strongest evidence for a negative relationship between ethnic diversity/divisions and public goods provision is Alesina *et al.* (1999), which draws on U.S. census data and budget information from all cities, metropolitan areas, and urban counties with at least 25,000 people to consider a range of dependent variables on spending. Consistent with the diversity debit hypothesis, the article shows that ethnic fractionalization (measured in racial terms) is negatively associated, and statistically significant, with the share of public spending on roads, education, welfare, and sewage and trash pickup. While this analysis is notable in the strength and detail of the budgetary data employed, there are several problems with the wide application of these findings. First, as the authors themselves take care to highlight, some of their results – including those on spending on health – are *not* consistent with a negative relationship.[6] Second, a subsequent examination of the same data (Gisselquist, 2014) finds that results are weakened when controls for state-level effects are included. As fiscal responsibilities and regulations differ across states, as does ethnic fractionalization, failing to control for state-level effects has arguably biased Alesina *et al.*'s results on fractionalization. Gisselquist (2014) also argues for the relevance of examining additional outcomes, showing in particular a positive relationship between fractionalization and levels of educational expenditure per child, which is a better measure of public goods provision than the share of total public spending on education.

Finally, to the best of our knowledge no similar findings have been replicated outside the U.S., and particularly not in sub-Saharan Africa where the hypothesis is so often applied. Even if we accept that there is a generally negative relationship between ethnic divisions and spending on public goods in the U.S., it is plausible that the results of this study are not externally valid because ethnic relations and governance are comparatively unique in the U.S.[7] With respect to sub-Saharan Africa in particular, perhaps the most obvious distinction concerns the relevance of formal *versus* informal institutions in governance and the ways in which governance functions differently in relatively high rule of law settings, such as the U.S., as compared to lower rule of law settings and neopatrimonial regimes, as in much of Africa (Bratton & van de Walle, 1994). For Zambia in particular, von Soest (2007) finds that informal, not formal institutions, are key to understanding state resources and tax collection. In situations where informal institutions are more important than formal institutions in policy-making, predictions drawn from formal models of voting (as in Alesina *et al.*) can be particularly problematic.

That said, a few notable studies conducted at the sub-national level in Africa do provide evidence consistent with diversity debit predictions; however, they have focused on the relationship between ethnic divisions and a more limited range of public goods outcomes than Alesina *et al.*, including education (funding, quality of facilities, and textbook ownership) and the maintenance of water wells, in non-representative population samples of Kenya and Tanzania (Miguel, 2004; Miguel & Gugerty, 2005).

In short, and contrary to the conventional wisdom, the extant literature shows that empirical evidence for the diversity debit hypothesis is by no means conclusive at the sub-national level, with some studies showing mixed or limited support and others suggesting the opposite relationship. The extant literature thus suggests the value of further empirical work on whether such a relationship exists (1) at the sub-national level, (2) in neopatrimonial African countries, and (3) for a wider and more disaggregated range of public goods outcomes. We turn to such analysis in the next section of this article. While our focus is on empirical testing of whether diversity debit predictions hold, we return at the end of this article to possible explanations of the discordant empirical findings thus established.

Before turning to the Zambian case, it is worth clarifying what we would expect to find if the diversity debit hypothesis were correct. The basic prediction is that a negative relationship should be displayed between diversity and various public goods outcomes. In light of the discussion above regarding the different roles and responsibilities of national *versus* sub-national governments, however, we focus in our analysis on a simple refinement of the basic prediction. For countries in which funding for public goods like health and education is highly centralized – as in much of sub-Saharan Africa, including Zambia – we amend the basic prediction as follows: First, we distinguish between budgetary outcomes and other measures of government performance, in particular related welfare outcomes. The diversity debit theory predicts a negative relationship between ethnic diversity and expenditure on public goods. In the fiscally centralized case, however, local governments and constituencies should have, by definition, little direct, formal influence on budgets as budgetary allocations and policies are decided at the center.[8] Thus, *if the diversity debit hypothesis operates through one or more of the direct local channels outlined above*, the characteristics of local communities should not directly influence total expenditure (including central government transfers), but they should influence implementation, how communities use and interact with government services, and related welfare outcomes. For instance, diversity's relation to preferences or social capital may influence parents' decisions to enroll their children in school, and thus educational enrollment rates. *Our revised prediction therefore is: If the diversity debit hypothesis is correct, and if it operates through one or more of the direct local channels outlined above, we should find a negative relationship between diversity and welfare outcomes, but we would not expect to find a relationship between diversity and total budgetary outcomes.*

This revised prediction implies that the relationship between diversity and public goods at sub-national levels may differ for different public goods outcomes and that whether there is a relationship with some outcomes – namely, budgetary ones – might depend on the degree of fiscal decentralization. It further underscores that the standard arguments underlying the diversity debit hypothesis as outlined above may fail to predict outcomes at sub-national levels in particular because they do not take into account the influence of interaction between levels of government. In other words, they are "single-level" theories focusing on the channels through which, for instance, local-level characteristics directly impact decisions and actions within local entities that influence outcomes measured at the local level – when it is plausible that local-level characteristics influence decisions and actions taken at the *national*-level, which in turn influence outcomes measured at the local level. This latter multi-level process would be consistent with the literature on distributive politics exploring how politicians target public expenditure to certain regions in order to support either core or swing voters (see Cox & McCubbins, 1986; Dixit & Londregan, 1996; Golden & Min, 2013; Kramon & Posner, 2013; Lindbeck & Weibull, 1987).

In contexts where political support and voting are associated with ethnicity, furthermore, we might also find a relationship between sub-national ethnic *characteristics* – not necessarily diversity – and sub-national public goods provision, including budgetary outcomes. For instance, if politicians provide more resources to swing constituencies (over core constituencies), and political support bases are ethnic, public spending may be higher in constituencies that are either ethnically-mixed *or* consist mainly of members of a "swing" ethnic group (Wantchekon, 2003). In a situation where swing constituencies tend to be ethnically-mixed, a positive relationship between diversity and budgetary outcomes would then be observed. On the other hand, if politicians support core ethnic constituencies instead, the relationship should be between ethnic strongholds and higher spending.

The literature on distributive politics thus suggests some explanatory hypotheses to be developed in future work. We consider several controls for political competition in our analysis in an exploratory way, and return to related hypotheses in Section 6 at the end of this article.

Finally, Gibson and Hoffman (2013) provide a useful point of departure for our analysis of Zambia as it speaks precisely to the relationship between sub-national ethnic diversity and government expenditure. Their finding of a positive relationship between diversity and local spending clearly runs counter to the basic diversity debit prediction. However, as they exclude central government transfers, their analysis cannot be explained by the basic distributive politics story outlined above and focuses only on a negligible share of public spending. This underscores the importance of a deeper examination.

Despite continuing efforts toward decentralization, government expenditure in Zambia – as in most sub-Saharan African countries – remains highly centralized. Most social sector spending on health and education is channeled through the central government via de-concentrated units at district level (District Health Management Teams – DHMTs, and District Education Boards – DEBs). In fact, total local council spending accounts for less than 5% of Zambian general government expenditure (Republic of Zambia, 2008). Moreover, local spending on service delivery is even smaller. In 2007, for instance, local councils spent on average a mere 10% of their expenditure on service delivery (Leiderer *et al*., 2012). Thus, to understand how diversity relates more broadly to public goods outcomes in Zambia, it is necessary to consider central government transfers rather than local-level expenditure, along with other measures of implementation and related welfare outcomes. We present such an analysis in Section 4.

3. DATA

We use a new purpose-built disaggregated dataset for Zambia covering the period 2004–09 at district level.[9] To address biases that may be caused by the fact that cities tend to be both more ethnically diverse and wealthier than rural and peri-urban areas, we exclude the four cities in the country from our analysis and limit our sample to 68 districts (14 municipal and 54 rural).

Data are drawn from the following sources: the Census of Population and Housing for 2000 and 2010 for information on ethnicity and language use; the 2006 Living Conditions Monitoring Survey (LCMS) for information on poverty levels and other district characteristics; the annual Government Financial Reports (GFR) for the years 2004–09 for information on budget allocations from the central government to districts; and administrative data provided by the Ministry of Health and the Ministry of Education for health and education outcomes. Most of these sources are generally not readily available and were obtained mostly from government sources by one of the authors during extensive field work in Zambia during 2009–11. While it is important to recognize the potential weaknesses and inaccuracies in such official statistics and financial reports, which have been well-documented in the African context (Anderson & Cheeseman, 2013; Jerven, 2013; Reinikka & Svensson, 2004; van de Walle, 2001), these sources provide the best data on these topics currently available for Zambia.

This new dataset offers important benefits over data used in previous studies of ethnic diversity and public goods provision. First, it allows us to examine the relationship between ethnic diversity and executed government expenditure to districts and to control for central government expenditure in analysis of other district-level governance outcomes. Second, and unlike most previous studies which have been conducted for a single point in time only, the data allow us to examine longer term behaviors in a dynamic setting.

(a) *Measuring diversity*

In Zambia, both tribal and linguistic cleavages have been salient in politics and governance (Lindemann, 2011a, 2011b; Posner, 2005). The 2000 Census includes information on individuals' self-reported ethnicity as well as on each person's predominant language of communication. Both variables are coded in the census according to the same 61 local languages, grouped into seven main ethno-linguistic groups: Bemba, Tonga, North-Western, Barotse, Nyanja, Mambwe, and Tumbuka; plus an eighth group for "other", which for the predominant language-use variable includes English (Republic of Zambia, 2000, p. 101).[10]

For the purpose of this article, we capture ethnic diversity in terms of the index of ethno-linguistic fractionalization (ELF) at the district level using both "ethnic" and "language" categories. While the literature offers a number of alternative measures, we use this measure because it is in the widest use by far, including in the work cited above (McDoom & Gisselquist, 2015). Fractionalization is calculated as:

$$ELF_i = 1 - \sum_{g=1}^{n} q_{gi}^2 \tag{1}$$

where q_{gi} is the population share of language or ethnic group $g = 1 \ldots n$, in district i. The obtained ethnic (ELF-E) and linguistic (ELF-L) fractionalization indices are highly but not perfectly correlated, with a correlation coefficient $r = 0.85$.[11] As a robustness check, we estimate the models presented in Section 4 with both indices.

(b) *Measuring public goods provision via government expenditure*

Zambia introduced government-wide activity-based budgeting in 2004 and government financial reports (GFRs) follow a mixed administrative and program classification. This allowed us to extract information on central government expenditure on health and education in each district.

Education expenditure by district is recorded in the GFRs under two different budget lines, "Regional Headquarters" and "Basic Schools." Under the former, expenditures administered by each District Education Board (DEB) are recorded, including grants for free basic education and infrastructure development. Under the latter, two expenditure items are reported over which the local DEBs formally have no influence: salaries and other emoluments, and grants to basic schools that are transferred from the Ministry of Education to DEB offices from where they are distributed among schools according to an allocation formula based on school characteristics such as enrollment figures, number of classes, and a gender parity factor (IOB, 2008, p. 31; World Bank, 2008, p. 76).[12] For health, all expenditure at district level is recorded under the budget line of the respective District Health Management Team (DHMT).

For the purpose of this study, we extracted a total of seven expenditure categories at district level for the years 2004–09. For education, these include total education expenditure, DEB administered expenditure, basic schools allocations, grants to basic schools, and teachers' salaries. For health, these are total health expenditure and expenditure on health service delivery (which is available only for 2006–09 and includes various sub-items; see Table 1).

In addition to total allocations, in each category, we calculate annual per capita expenditure using (interpolated) district population figures taken from the 2000 and 2010 Census. For the teachers' salaries and grants for basic schools, we use the population in the relevant age group of primary school pupils (7–13 years) in 2000 to calculate per capita figures.[13]

For each budget item, the GFR reports budget estimates, authorized provisions, actual expenditure, and budget variance. In addition to actual expenditure, we are thus also able to calculate budget execution rates i.e., the ratio of money expended compared to releases received by the respective spending unit for each expenditure item, which provide a measure of the deconcentrated government units' operational efficiency or absorptive capacity.

(c) *Assessing public goods provision via education and health outcomes*

Data on education outcomes are relatively limited, covering the number of pupils in grades 1–7 and grades 1–9 for the period 2004–09. From this and using census data on population by age group, we were able to construct gross enrollment rates for primary education (grades 1–7) and lower secondary education (grades 8 and 9). In addition, we obtained data on the total number of pupils and teachers (grades 1–12) for 2008; and the number of schools, pupils, and teachers in basic education for 2009.

Data on health are more comprehensive, although they cover a shorter time frame (2004–08). Based on the available data, we calculated 11 health indicators that include: the number of beds in health facilities per 1,000 inhabitants; health center staff per capita; hospital outpatient department staff per capita; the maternal mortality rate; the under-five mortality rate; the rate of underweight children under five; under one year olds' immunization rates for tuberculosis, diph theria–pertussis–tetanus, and polio; and the rate of fully immunized children under one year.

(d) *District-level covariates*

District-level covariates that are expected to affect government expenditure and/or welfare outcomes were constructed using information from various sources. As a general measure of the level of deprivation, we calculated district-level poverty headcounts using data from the 2006 LCMS. Annual district-level population estimates were calculated by interpolating data from the 2000 and 2010 Population Census. Data on the district surface area measured in square kilometers were extracted from the 2000 Population Census to capture the spatial dimension of districts.

A dummy variable was constructed using the 2006 LCMS, to control for possible infrastructure and scale economy effects from rural environments. We also used the routing function of the Google Maps™ mapping service to construct a variable that measures the distance by road from each district capital to the national capital Lusaka, as a measure of geographic location and remoteness.

Finally, and in order to control for possible political targeting of social sector spending, we used information on election results in national presidential elections in 2001 and 2006 from the Zambia Electoral Commission to construct a variable that calculates the district share of votes for the ruling party, Movement for Multiparty Democracy (MMD) in the 2001 and 2006 presidential elections, using 2001 values for all years up to 2006 and 2006 values for 2007 and subsequent years. As a measure of local political competition, we calculated the

Table 1. *Summary statistics for rural and municipal districts, 2004–09*

Variable	Definition	Data source	*N*	Mean	SD	Min	Max
Ethnic fractionalization by district							
ELF-E	ELF measure based on main "ethnic" groups as identified in the Census	Census 2000	68	0.307	0.227	0.013	0.777
ELF-L	ELF measure based on main "language" groups as identified in the Census	Census 2000	68	0.201	0.182	0.005	0.683
District-level covariates							
Area	Surface area (km^2)	Census 2000	68	11	8	1	41
Poverty	Poverty headcount in 2006	LCMS 2006	68	0.695	0.161	0.224	0.961
Population	Logarithm of population	Census 2000/10	408	11.694	0.554	9.975	13.003
Distance Lusaka	Distance from the district capital to Lusaka (km)	Google Maps™	68	577	294	45	1170
Rural	Dummy variable = 1 if district is classified as rural (*vs.* municipal or city)	LCMS 2006	68	0.794	0.405	0	1
Vote MMD	Vote share for the ruling party in the most recent presidential election prior to the year of analysis (2001 for 2001–06 and 2006 for 2007–09)	ECZ	136	0.418	0.200	.077	0884
Share MMD councilors	% share of district council members belonging to the ruling party (2006)	ECZ	65	56.047	31.013	0.000	100
Share Bemba	% of population self-identifying as belonging to the Bemba ethnic group	Census 2000	68	34.062	39.288	0.214	99.324
Share Tonga	% of population self-identifying as belonging to the Tonga ethnic group	Census 2000	68	16.876	30.511	0.055	93.394
Share North-West group	% of population self-identifying as belonging to the North-West ethnic group	Census 2000	68	14.666	27.988	0.067	98.371
Share Barotse	% of population self-identifying as belonging to the Barotse ethnic group	Census 2000	68	9.239	21.560	0.021	87.619
Share Nyanja	% of population self-identifying as belonging to the Nyanja ethnic group	Census 2000	68	12.996	27.811	0.133	96.966
Share Mambwe	% of population self-identifying as belonging to the Mambwe ethnic group	Census 2000	68	6.402	18.176	0.014	88.917
Share Tumbuka	% of population e self-identifying as belonging to the Tumbuka ethnic group	Census 2000	68	4.019	13.722	0.022	88.189
Central government district expenditure (in billion kwacha)							
Education expenditure	Total central government expenditure for education in each district	GFR	408	10.537	6.373	0.809	47.020
Education expenditure p.c.	Education expenditure per capita	GFR	408	0.092	0.057	0.006	0.322
DEB expenditure	Total central government expenditure to DEB	GFR	408	0.639	1.186	0.046	16.430
DEB expenditure p.c.	DEB expenditure per capita	GFR	408	0.006	0.007	0.000	0.088
Basic school expenditure	Total central gvt. expenditure for basic schools (incl. personal emoluments and grants)	GFR	408	9.989	6.315	0.457	46.662
Basic expenditure p.c.	Basic schools expenditure per capita	GFR	408	0.087	0.055	0.004	0.310
Grants basic schools	Expenditure on grants to basic schools	GFR	406	0.875	1.105	0.000	5.675
Grants basic schools p.c.	Expenditure on grants to basic schools per capita	GFR	406	0.031	0.041	0.000	0.252
Teachers' salaries	Expenditure on teachers' salaries	GFR	400	5.821	4.375	0.013	37.116
Teachers' salaries p.c.	Expenditure on teachers' salaries per capita	GFR	400	0.198	0.124	0.001	1.020
Health expenditure	Total DHMT team expenditure	GFR	408	4.193	4.986	0.014	33.469
Health expenditure p.c.	Total DHMT expenditure per capita	GFR	408	0.030	0.034	0.000	0.450
Health service delivery	Health service delivery expenditure; included items are: epidemic preparedness, provision of 1st level referral services, roll back malaria, HIV/AIDS/STIs, tuberculosis, integrated reproductive health, child health, environmental health, mental health, oral health	GFR	272	1.403	0.981	0.062	7.812
Health service delivery p.c.	Health service delivery expenditure per capita	GFR	272	0.011	0.005	0.001	0.041
Budget execution rates							
Education execution	Education budget execution rate	GFR	408	0.993	0.085	0.516	1.264
DEB execution	DEB budget execution rate	GFR	408	0.836	0.180	0.251	1.126
Basic schools execution	Basic schools budget execution rate	GFR	408	1.005	0.107	0.378	1.392
Basic school grant execution	Grants to basic schools budget execution rate	GFR	406	1.391	1.224	0.000	6.308
Teachers' salaries execution	Teachers' salaries budget execution rate	GFR	400	0.993	0.040	0.500	1.000
Health execution	DHMT budget execution rate	GFR	408	0.670	0.293	0.010	1.302
Health service execution	Health service delivery budget execution rate	GFR	272	0.813	0.242	0.098	2.137
Educational outcomes							
Prim. school enrollment	Primary school enrolment (grades 1–7)	MoE	408	1.234	0.247	0.577	2.047
Low sec. enrollment	Lower secondary school enrollment (grades 8 and 9)	MoE	408	0.552	0.269	0.086	1.541
No. schools 2008	Number of schools (all schools) in 2008	MoE	68	116	52	21	278
No. teachers 2008	Number of teachers (all schools) in 2008	MoE	68	910	526	213	2799
No. basic schools 2009	Number of basic schools in 2009	MoE	68	114	52	18	285
TPR 2008	Teacher–pupil ratio (all schools) in 2008	MoE	68	0.022	0.005	0.013	0.038
Basic school TPR 2009	Teacher–pupil ratio for basic schools in 2009	MoE	68	0.019	0.005	0.009	0.035
Health outcomes							
Total beds	Number of beds in health facilities per 1,000 population	MoH	329	2.236	1.073	0.705	6.531
HC staff	Health center staff per 10,000 population	MoH	329	4.208	2.128	0.687	12.305
Hospital OPD staff	Hospital outpatient department staff per 10,000 population	MoH	255	0.704	0.611	0.000	3.836
BCG immunization	Under 1 year olds' immunization rate for tuberculosis	MoH	329	1.223	0.225	0.432	2.442
DPT3 immunization	Rate of under 1 year olds with three doses of the combined diphtheria/pertussis/tetanus vaccine	MoH	329	1.118	0.239	0.491	2.131
OPV3 immunization	Rate of under 1 year olds with three doses of oral polio virus vaccine	MoH	329	1.095	0.260	0.488	2.437
Measles immunization	Under 1 year olds' immunization rate for measles	MoH	329	1.017	0.202	0.481	2.123
FIC immunization	Rate of fully immunized under 1 year old children	MoH	329	0.821	0.177	0.312	1.806
Maternal mortality	Maternal mortality (log maternal deaths per 100,000 live births)	MoH	289	5.354	0.743	2.954	7.437
Under 5 mortality	Under five mortality (deaths of children under 5 per 1,000 live births)	MoH	328	57.001	30.755	2.542	221.24
Underweight under 5	Underweight children under the age of 5 per 100 under 5 year olds weighed	MoH	329	13.856	7.426	0.945	35.857

Note: GFR: Government Financial Reports ("Blue Books"); LCMS: Living Conditions Monitoring Survey; DEB: District Education Board; DHMT: District Health Management Team; ECZ: Electoral Commission of Zambia; MoE: Ministry of Education; MoH: Ministry of Health.

share of MMD councilors on the district council after the 2006 elections. We also consider the population share of each main ethnic group, along with fractionalization, as a possible control for ethnically-based political targeting (see Table 1).

An initial examination of the level of association between the index of ethnic fractionalization (ELF-E) and most measures of public goods provision, including education and health outcomes, reveals a *positive* and statistically significant association at conventional levels (see Table 2). The correlation coefficients vary in terms of magnitude but overall they seem to favor a diversity dividend hypothesis. In the next section we undertake a more rigorous econometric analysis to examine this relationship.

4. MODEL SPECIFICATION AND ECONOMETRIC METHODS

Based on the theoretical discussion presented in Section 2, we undertake the empirical analysis in two steps. First, we begin in Section 4(a) with a model that tests the diversity debit hypothesis with regard to spending on public goods. As above, in centralized countries such as Zambia we predict no relationship based on the major mechanisms elaborated in the literature on the diversity debit hypothesis. As a second step, we develop a model in Section 4(b) that tests the diversity debit hypothesis with regard to welfare outcomes. As above, in centralized countries such as Zambia, we predict that a negative relationship should be evident here if the diversity debit hypothesis is correct.

(a) *The government expenditure model*

More formally, we derive a model that takes the form:

$$s_{it} = \alpha_{it} + \beta x_{it} + \lambda f_i + \mu_i + \zeta_t + v_{it} \qquad (2)$$

where the subscripts i and t denote district and year, respectively; s_{it} measures various items of government expenditure on education or health; x_{it} is a vector of district-level covariates that are expected to affect the government's allocation decisions, including (i) the logarithm of district population to control for scale effects with respect to central government allocations; (ii) the local poverty headcount index as a measure of deprivation that may capture the existing demands for social services at local level; (iii) the district surface area to capture the spatial dimension of districts that may affect the transaction costs associated with public goods provision; (iv) a binary indicator that identifies rural communities to capture possible infrastructure and scale economy effects; (v) the distance to the national capital Lusaka as a measure for remoteness and access to the center of political power, and (vi) the ruling party's vote share in past presidential elections, which controls for the possibility of political targeting of health and education spending. f_i is the ethnic or linguistic fractionalization index, the parameter of which measures the effect of ethnic diversity on government expenditure; μ_i denotes unobserved district-specific and time-invariant effects; ζ_t is a vector of time dummies capturing universal time trends, whereas α, β, λ, and v are the intercept, the parameter estimates and the idiosyncratic error term, respectively. We estimate the model for each of the expenditure items both for levels of total expenditure as well as in per capita terms.

The basic prediction of the diversity debit hypothesis can be generally understood to be that the parameter of interest, λ, is negative and statistically significant. However, as above, our expectation at the sub-national level in centralized countries is for no relationship.

There are some important constraints with regard to estimating the effect of ethnic diversity on public expenditure as formulated by Eqn. (2). Ideally we would want to exploit the within-district variation to estimate Eqn. (2) using fixed-effects estimates in order to control for any unobserved district-level characteristics that may affect central government allocation decisions. However, while we observe variation in the expenditure variables over time, the fractionalization indices as well as most covariates including poverty, district surface area, distance to Lusaka, and the rural dummy are time-invariant.

Furthermore, budget allocations tend to be path dependent, with annual budget plans usually building incrementally on allocations in preceding fiscal years. Incremental budgeting implies that the expected errors are likely to be serially correlated over time.

Given these data constraints we resort to estimate Eqn. (2) using a panel feasible GLS estimator that corrects for first-order autocorrelation within panels. As a robustness check, we also estimate Eqn. (2) with (i) a pooled ordinary least squares (OLS) estimator with standard errors corrected for correlation across panels, which allows for different structures of the error term, and (ii) a pooled OLS estimator that assumes correlation across panels and more general serial correlation in the error, following the method of Driscoll and Kraay (1998).

(b) *The welfare outcomes model*

Regarding the relationship between ethnic diversity and welfare outcomes, we estimate two models. The first model measures the effect of ethnic diversity on a number of education or health outcomes, after controlling for the district-level covariates included in Eqn. (2), and the effect of central government expenditure on education or health, respectively, which captures the government's decisions to allocate public funds to the districts, regardless of the size of the local population. The second model measures the effect of ethnic diversity on the same welfare outcomes, after controlling for the same district-level covariates and the effect of per capita

Table 2. *Correlations between ethnic fractionalization and public good provision measures*

	ELF-E
ELF-E	1
Primary school enrollment	0.25***
Lower secondary enrollment	0.31***
Number of teachers (2008)	0.41***
Teacher–pupil ratio (2008)	0.3**
Maternal mortality	−0.17***
Under 5 mortality	−0.21***
Underweight under 5	−0.49***
Total beds p.c.	−0.16***
Health center beds p.c.	−0.3***
Hospital beds p.c.	−0.13**
Health center staff p.c.	0.47***
Hospital OPD staff p.c.	0.29***
BCG immunization	0.05
DPT3 immunization	0.13**
OPV3 immunization	0.14**
Measles immunization	0.12**
FIC immunization	0.44***

Note: coefficients at conventional levels; *10% significance level; **5% significance level; ***1% significance level.

expenditure, which now accounts for the effect of resource distribution across the local populations. More formally, the outcome equations take the following form:

$$w_{it} = \alpha_{it} + \beta x_{it} + \phi s_{it} + \lambda f_i + \mu_i + \zeta_t + e_{it} \quad (3)$$

where, as before, the subscripts i and t denote district and year respectively; w_{it} measures the education and health outcomes; x_{it} is the vector of district-level covariates derived in (2) that are expected to affect the welfare outcomes; s_{it} measures total (or per capita) government expenditure on education or health; and f_i again measures ethnic or linguistic fractionalization. μ_i, ζ_t, α, β, ϕ, and λ, are as defined above, whereas e is the idiosyncratic error term.

The diversity debit hypothesis would predict the parameter of interest, λ to be negative and statistically significant. We note, however, that government expenditure is likely to be endogenous. It is reasonable to expect that welfare outcomes at district level are influenced by the allocation of public resources, as much as the decision on how to distribute such resources is likely to be influenced by local demands and social needs. The presence of endogeneity would imply that s_{it} is correlated with e_{it}, and therefore under an OLS framework, Eqn. (2) would yield biased and inconsistent estimates. To test and address the endogeneity problem, we resort to instrumental variable estimators, including two-stage least squares (2SLS), limited information maximum likelihood (LIML), and generalized method of moments (GMM) to obtain, under a pooled cross-sectional setting, the following system of equations:

$$s_{it} = \alpha_{it} + \beta x_{it} + \lambda f_i + \delta \mathbf{z}_{it} + \mu_i + \zeta_t + \upsilon_{it} \quad (4)$$

$$w_{it} = \alpha_{it} + \beta x_{it} + \phi \hat{s}_{it} + \lambda f_i + \mu_i + \zeta_t + \upsilon_{it} \quad (5)$$

where $\mathbf{z}_{it}$ is a vector of strictly exogenous instrumental variables that are partially correlated with s_{it}, so the coefficient of $\mathbf{z}_{it}$ is nonzero, i.e., $\delta \neq 0$ and $Cov(\mathbf{z}_{it}, \upsilon_{it}) \neq 0$, while $\mathbf{z}_{it}$ is uncorrelated with w_{it}, so $Cov(\mathbf{z}_{it}, e_{it}) = 0$. Finding valid instruments is thus important. We experiment with two approaches: First, we exploit exogenous instrumental variables that have been used previously in the literature. Specifically, we use the logarithm of population, and the distance to the national capital, Lusaka, as external instruments. With regard to the former instrument, Easterly and Rebelo (1993) and Gebregziabher and Niño-Zarazúa (2014) find that the scale of the economy, measured by its population, is an important determinant of fiscal policy in general, and the allocation of social expenditure in particular. Larger populations would have the effect of diminishing the average cost of providing public goods. These scale effects would arise from the nonconvexities associated with the costs of public goods provision. Yet, there is no reason to suspect that a particular district will achieve higher or lower levels of welfare simply because it has more or less people. The second instrument is based on the observation made in previous studies that more remote areas in Zambia tend to receive lower transfers from the central governments. Picazo and Zhao (2009), for instance, find that the most remote and least urbanized areas in Zambia receive the lowest per capita releases in the health sector (De Kemp, Faust, & Leiderer, 2011). This could imply that there is a negative correlation between the distance to the capital city and the bargaining power that rural communities are able to exercise to attract public resources from the center; or that remote districts are sanctioned more frequently in financial terms if they fail to meet formal planning or reporting requirements. It is not entirely clear, however, whether more or less financial resources would necessarily lead to better or worse welfare outcomes, after controlling for poverty and the rural environment. We favor the use of log of population as our preferred instrument, given past evidence about its validity, but also experiment with distance to Lusaka as a secondary instrument.

An initial examination of the pairwise correlations between government expenditure, both in total and in per capita terms, and the identified instruments show high correlations between the endogenous variables and the log of population, with most coefficients exceeding $r = 0.45$ values; however, the correlations become moderately low when using the distance to the capital city as instrument, with r values ranging from 0.12 to 0.15.

To verify the validity of the instruments, we follow Stock and Yogo (2005) to test for the concern of weak instruments that can lead to size distortions of the Wald test on the parameters. The results for education outcomes show that the Eigenvalue statistic and the F statistic comfortably exceed the critical values of the Stock and Yogo statistic at 5% or 10% for both the 2SLS and LIML models, which allows us to reject the null hypothesis of weak instruments, particularly for the case of log of population, but also when combined with distance to the capital city. While distance to the capital city alone appears as a weaker instrument than log of population when running the models of education outcomes, it becomes stronger when running the models of health outcomes, especially when health expenditure is instrumented in per capita terms (see Table 3). Therefore, we present the results in Section 5 and also in the Appendix using the individual and combined instrument sets.

Given the validity of the instruments, we resort to the Hausman procedure (Hausman, 1978) to test for the assumption of endogeneity of government expenditure using Eqns. (3) and (5) so $TH = (\hat{\phi}_{2SLS} - \hat{\phi}_{OLS})^2 / \hat{V}(\hat{\phi}_{2SLS} - \hat{\phi}_{OLS})$ is $\chi^2(1)$ distributed under the null of exogeneity. As a robustness check, we also compute the Durbin–Wu–Hausman (DWH) test, which in addition produces robust test statistics (Davidson, 2000). The Hausman and DWH results for government expenditure on education and health strongly reject the null of exogeneity for most outcome variables (see Tables in Appendix A). Therefore we conclude that government expenditure is endogenous and thus favor the use of instrumental variables estimators over OLS in the analysis.

Since we have longitudinal data, with most of the education and health outcomes being observed over the 2004–09 period, we extend the analysis to system-GMM (SGMM) estimators in a dynamic setting, exploiting both the internally generated instruments, and also their combination with the external instruments. This strategy allows us to verify the robustness of our results, in the absence of external instruments. Under a dynamic framework, Eqn. (5) can be rewritten as follows:

$$w_{it} = \alpha + \theta w_{it-1} + \beta x_{it} + \phi \hat{s}_{it} + \lambda f_i + \mu_i + \zeta_t + \upsilon_{it} \quad (6)$$

where w_{it-1} and θ are the lag of the dependent variable and its parameter estimate, respectively. The presence of district fixed-effects, μ_i, would suggest that the preferred approach is a fixed-effects model, which would allow to mitigate the heterogeneity-induced bias and control for district-related endogeneity. However, the inclusion of lagged dependent variables would produce inconsistent fixed-effects estimates.

The SGMM estimator developed by Blundell and Bond (1998) circumvents the endogeneity problem by solving a system of level and difference equations. Lagged differences of the endogenous variables are used as instruments in the level equations, while lagged levels of the endogenous variables are used as instruments in the first differenced equations. SGMM improves the accuracy of estimates by exploiting

Table 3. *Stock-Yogo test for weak instruments*

			Instruments		
Instrumented variable			Log population + km to Lusaka	Log population	Km to Lusaka
Education expenditure	*F*-Statistic		40.67	78.96	15.12
	Minimum eigenvalue		57.74	107.1	10.28
Education expenditure per capita	*F*-Statistic		77.46	145.6	3.947
	Minimum eigenvalue		154.4	277.5	4.111
Health expenditure	*F*-Statistic		54.06	93.32	19.28
	Minimum eigenvalue		51.20	78.95	23.62
Health expenditure per capita	*F*-Statistic		6.37	7.27	8.00
	Minimum eigenvalue		10.62	3.79	15.43
		10%	15%	20%	25%
Critical values 2SLS/LIML size of nominal 5% Wald Test		16.38	8.96	6.66	5.53

additional moment conditions that are informative in the presence of persistent data. Hence, we opt for a SGMM estimator with external instruments as our preferred model in Section 5(b), the robustness of which we test by comparing the results with the SGMM using the internally generated set of instruments, and also the 2SLS, GMM and LIML models. [14]

5. RESULTS

(a) *On ethnic diversity and budgetary outcomes*

Contrary to our revised prediction, but consistent with the basic prediction of the diversity debit hypothesis, we find a clearly negative relationship between ethno-linguistic fractionalization and central government expenditure at district level in both sectors and across budget lines. Table 4 shows a summary of the panel regression results for each expenditure item in absolute as well as per capita terms for both fractionalization indices ELF-E and ELF-L. [15] Besides ethno-linguistic fractionalization, the log of district population and the distance by road to Lusaka appear to be good predictors of differences in budget allocations between districts. As expected, the log of population has a highly significant positive coefficient in all specifications with total budget allocations, and a negative one for per capita allocations, except for per capita health spending. The distance to Lusaka is significant and negative throughout for total and per capita expenditure, except for grants to basic schools and health service delivery, where it is significant only in some model specifications. [16]

In contrast, and different from recent studies on fertilizer subsidies (Mason, Jayne, & van de Walle, 2013) or infrastructure projects (Leiderer, 2014) in Zambia, the results do not provide particularly strong evidence for political targeting of recurrent health and education expenditure, with the vote share received by the ruling MMD insignificant in all specifications except for DEB allocations and total and per capita allocations for health service delivery (positive) and grants to basic schools (negative). [17] The same goes for the share of MMD members in the district council, as well as the shares of the main ethnic groups in the population: including these variables yields mostly insignificant coefficients and, more importantly, does not alter the results with respect to the negative relation between ethno-linguistic fractionalization and expenditure. [18]

Likewise, the poverty headcount is insignificant in most specifications, except for per capita grants to basic schools and total health allocations, and health service allocations, for which it has a (weakly significant) positive coefficient. This suggests that social sector expenditure was not markedly "pro-poor" in Zambia during the second half of the past decade.

As described above, we run as robustness checks the same equations with different specifications that allow for correlation across districts and autocorrelation of up to four lags. The alternative specifications produce very similar results with consistently smaller standard errors. [19]

The results from the model described in Eqn. (2) may seem to confirm the diversity debit hypothesis that suggests a negative effect of local ethno-linguistic fractionalization on public goods provision via central government spending. However, it is *prima facie* not clear, by which diversity debit mechanism local diversity should affect central government's spending decisions in a highly centralized governance system such as Zambia's, in particular so as measures of political competition or ruling party dominance do not seem to play a role in central government allocations on recurrent expenditure.

One explanation for the observed pattern could be that it is not budget allocations, but the absorptive capacity of local districts that differs between districts with varying degrees of ethnic fractionalization. In Zambia, local offices of the central government ministries act as the central government's spending units at district level. The amount of money spent on public goods depends not only on the amount of resources the central government allocates to them, but also on the effectiveness and efficiency with which these spending units make use of the available resources. If the hypothesis that ethnic diversity leads to less efficient institutions and governance is correct, then we might expect the absorptive capacity of local spending units to be negatively correlated with ethnic diversity.

To control for this possibility, we estimate Eqn. (2) with budget estimates as well as execution rates (i.e., the ratio between releases to each district education board, and district health management team, and their corresponding executed expenditure) as dependent variables. [20]

However, the findings, which are presented in Table 1 of Appendix C, do not support the absorptive capacity hypothesis. The results for budget estimates are strongly in line with those for actual expenditure figures, whereas the estimates for budget execution rates are insignificant throughout (Table 2 of Appendix C).

Taken together, our findings suggest that the observed negative relationship between ethnic fractionalization and central government spending is in fact due to allocation decisions taken at the central level and not because of differences in the absorptive capacity of the deconcentrated spending units

Table 4. *Panel GLS regression results for health and education expenditure*

Dependent variable	Main explanatory variable	
	ELF-E	ELF-L
Total education expenditure	–	–
Total education expenditure per capita	Negative*	–
DEB expenditure	Negative**	Negative**
DEB expenditure per capita	Negative*	Negative**
Basic schools expenditure	–	–
Basic schools expenditure per capita	Negative*	–
Grants to basic schools	–	–
Grants to basic schools per capita in relevant age group	–	–
Teachers' salaries	–	–
Teachers' salaries per capita in relevant age group	–	–
Health expenditure	Negative**	Negative*
Health expenditure per capita	–	–
Health service expenditure	Negative**	Negative**
Health service expenditure per capita	Negative**	Negative***

Note: – coefficient insignificant at conventional levels; *10% significance level; **5% significance level; ***1% significance level.

at district level. Yet, as noted in the previous section, if ethnic fractionalization has a direct effect on welfare outcomes at district level, then it is likely that these allocation decisions are endogeneous. In this case we would expect the estimates underlying Table 4 to be biased and inconsistent. The second step of our analysis addresses this constraint.

(b) *On ethnic diversity and welfare outcomes*

In the second stage of our analysis we follow Gerring *et al.* (2015), Miguel (2004), and Miguel and Gugerty (2005) in studying the direct link between ethnic diversity and welfare outcomes.

Table 5 presents the OLS, 2SLS, GMM, and LIML regression results for ethnic fractionalization and welfare outcomes, and the full regression results are reported in Appendix D. Column 1 shows the coefficients on ELF-E with significance levels for robust standard errors for the OLS estimator. Columns 2–3 show the 2SLS estimates for education and health expenditure instrumented with the log of district population (column 2), the distance to Lusaka (column 3), and both instruments (column 4). Columns 5 and 6 show the GMM and LIML estimates using both instruments.

For the education sector, the results show a clearly positive relationship across specifications between ethnic fractionalization and primary school enrollment, but none with the other outcome variables. For the health sector, the results are in line with those obtained from the education sector: there is a positive effect of ethnic diversity on all immunization rates, the under-five mortality rates, and the share of underweight children under five (where a negative sign means a reduction in underweight). The coefficients on maternal mortality also have the expected negative sign, but are statistically insignificant. The only exception is the number of beds in health facilities per 1,000 population, for which we find a significant negative relationship.

Given the specific setup of Zambia's health system, health indicators capture allocation decisions at different levels of government. Procurement of medical supplies, capital investment, and staff allocations are the responsibility of the Ministry of Health, whereas the DHMTs are responsible for service delivery at district level (ILO, 2008, p. 113). It is thus not surprising that the coefficient for beds in health facilities is negative, as it is most likely driven by the same decisions taken at the central government level that drive general health expenditure. The results for outcomes such as immunization are more likely to be driven by decisions at the local level, whereas staffing of health facilities is determined by central and local decision-making (Bossert, Chitah, & Bowser, 2003, p. 359).

The results are highly robust across the various model specifications and the selection of instruments for government expenditure. Moreover, the effects are comparable in terms of magnitude and direction across models. A one standard deviation increase in the ethnic fractionalization index leads to an increase of the primary school enrollment rate of between 5.7 and 6.9 percentage points. The effect of ethnic diversity on immunization rates is of comparable magnitude at between 4.1 percentage points (for BCG immunization rates in the GMM specification) and 8.2 percentage points (for fully immunized children in the 2SLS specification with only distance to Lusaka as instrument). The same increase in the ethnic fractionalization index reduces under-five child mortality by 4.1 and 5.1 per 1,000 live births and the number of underweight children per 100 weighed children under 5 by between 2.9 and 3.2.

Controlling for per capita rather than total expenditure does not alter the results for health substantially (see Table D1 in Appendix D), with the exception of individual vaccination rates in the 2SLS specification with only the (log) population instrument and the LIML model (except BCG immunization, which remains significant). For education, ELF-E becomes positive and significant for both lower secondary school enrollment and the number of schools in 2008 across all specifications (except number of schools 2008 in the 2SLS model with distance to Lusaka as the only instrument).

With most of the education and health outcomes being observed over various years, we extend the analysis to include SGMM estimators in a dynamic setting. Table 6 shows the SGMM estimates for the coefficient on ethnic fractionalization in different specifications with lagged dependent variables (columns 1–4) and without lagged dependent variables and the two external instruments, jointly and individually, inputted (columns 5–8). Individual regression tables for each outcome variable are reported in Appendix E.

The SGMM results show similar results in terms of direction, although for education outcomes the strength of the association is much weaker, with only primary school enrollment exhibiting a significant and positive effect in the

Table 5. *Overview of regressions results for ethnic fractionalization and social sector outcomes*

Method	OLS	2SLS[a]	2SLS[b]	2SLS[c]	GMM[c]	LIML[c]
Dependent variable				Coefficient on ELF-E		
Primary school enrollment	.261***	.305***	.249***	.300***	.304***	.303***
Lower secondary enrollment	.046	.090	−.002	.082	.037	.089
Teacher–pupil ratio (2008)	.001	.003	−.004	.002	.003	.003
Number of schools (2008)	−14.156	−40.042	−22.264	−38.088	−35.096	−38.644
Number of teachers (2008)	84.765	−62.003	−157.748	−72.526	−91.064	−74.558
Teacher–pupil ratio in basic schools (2009)	.003	.003	.003	.003	.003	.003
Number of basic schools (2009)	12.332	10.433	10.218	10.428	10.081	10.428
Total beds	−1.107***	−1.006***	−1.265**	−1.041***	−1.190***	−.974**
Health Centre Staff p.c.	2.554***	2.686***	2.170**	2.615***	2.413***	2.946**
Hospital OPD Staff p.c.	.284*	.350*	.218	.326*	.085	.334*
BCG immunization	.190***	.191***	.192***	.191***	.180***	.191***
DPT3 immunization	.240***	.241***	.244***	.242***	.230***	.242***
OPV3 immunization	.272***	.273***	.276***	.273***	.268***	.274***
Measles immunization	.205***	.206***	.206***	.206***	.206***	.206***
FIC immunization	.361***	.361***	.363***	.361***	.367***	.362***
Maternal mortality (log)	−.354	−.331	−.161	−.307	−.214	−.291
Under 5 mortality	−18.726**	−18.702**	−21.821**	−19.128**	−22.595***	−19.258**
Underweight under 5	−12.981***	−12.980***	−12.711***	−12.943***	−13.354***	−12.941***

Notes: Values show estimated coefficient for ethnic fractionalization index ELF-E. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ for robust standard errors.
[a] Instruments for central government expenditure using log population.
[b] Instruments for central government expenditure using distance from district capital to Lusaka.
[c] Instruments for central government expenditure using log population and distance from district capital to Lusaka.

specifications without the lagged dependent variable as an explanatory variable. For the health sector, however, the results remain fairly robust, with ethnic fractionalization having a strong and positive effect on immunization rates and a reduction in the number of underweight children under the age of five.[21]

Tables E16–E30 in Appendix E present the results from the SGMM equations instrumenting for per capita expenditure. With per capita expenditure, ethnic fractionalization becomes insignificant for primary school enrollment, whereas for health outcomes the results remain highly robust.

Running the SGMM models with ELF-L instead of ELF-E yields very similar results, with slightly larger coefficients for ELF-L than for ELF-E (see Table E32 in Appendix E).

Overall, we find strong evidence for a positive effect of ethno-linguistic fractionalization on health outcomes at district level in Zambia, particularly on immunization rates, and the number of underweight children. The results for the education sector are also consistent with our priors, although somewhat weaker than for the health outcomes.

Since we are calculating parameter estimates from a relatively large number of welfare outcomes, the probability of incurring in type I errors, i.e., the likelihood of making false discoveries, tends to increase with a larger number of confidence intervals from different coefficients. To address this problem, we control for the family-wise error rate (FWER), which is the probability that at least one true null hypothesis is rejected, using multi-test procedures, including the Bonferroni correction method, and

Table 6. *Overview results of SGMM estimation for ethnic fractionalization and all outcomes with total sector expenditure*

	With lagged dependent variable as regressor				Without lagged dependent variable as regressor			
Model	(1)[a]	(2)[b]	(3)[c]	(4)[d]	(5)[a]	(6)[b]	(7)[c]	(8)[d]
Dependent variable								
Prim. school enrollment	.033	.032	.037	.031	.250*	.253*	.263*	.230
Lower sec. enrollment	−.004	.007	.011	−.006	.001	.024	.024	−.016
Total beds	−1.091**	−.172	−.234	−.675	−1.450**	−1.384**	−1.477**	−1.653***
Health Centre staff p.c.	.524	−.016	.402	−.002	2.625**	3.058**	2.892**	3.096**
Hospital OPD staff p.c.	−.001	−.065	−.084	.001	−.098	.140	.166	−.086
BCG immunization	.116	.131	.120	.140*	.157*	.160	.152*	.180*
DPT3 immunization	.185**	.102*	.135**	.117*	.326***	.256**	.240**	.281**
OPV3 immunization	.244***	.151*	.193**	.151*	.305***	.289**	.243**	.301**
Measles immunization	.213**	.189**	.188**	.197**	.240**	.226**	.220**	.236*
FIC immunization	.309***	.201**	.309***	.189**	.356***	.422***	.362***	.415***
Maternal mortality (log)	−.028	.166	−.012	.203	−.158	.085	−.074	.081
Under 5 mortality	.772	3.049	1.259	1.730	−3.684	−8.747	−4.435	−8.052
Underweight under 5	2.989	.547	2.954	.844	−11.765***	−12.456***	−12.092***	−11.593***

Notes: Values show estimated coefficient for ethnic fractionalization index ELF-E. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ for robust standard errors.
[a] (1) and (5) use only internal instruments (second and longer lags of central government education/health expenditure).
[b] (2) and (6) use both internal and external (log population and distance from district capital to Lusaka) instruments.
[c] (3) and (7) use internal and one external (log population) instrument.
[d] (4) and (8) use internal and one external (distance to Lusaka) instrument.

Table 7. *Number of reject null hypotheses using different methods. Uncorrected p-value thresholds at 0.05 levels*

Models	FWER			FDR		
	Bonferroni	Holm (1979)	Holland and Copenhaver (1987)	Benjamini and Liu (1999a)	Benjamini and Liu (1999b)	Benjamini and Yekutieli (2001)
OLS	9	9	9	9	9	9
2SLS[a]	13	13	13	15	15	13
2SLS[b]	3	4	4	6	6	6
2SLS[c]	12	12	12	14	14	12
GMM[c]	12	14	14	18	18	14
LIML[c]	12	13	13	14	14	13
SGMM[d]	4	4	4	4	4	4
SGMM[e]	7	7	8	9	9	8

Note: a instruments for central government expenditure using log population; *b* instruments for central government expenditure using distance from district capital to Lusaka; *c* instruments for central government expenditure using log population and distance from district capital to Lusaka; *d* is estimated with lagged dependent variable and no external instruments; *e* is estimated with lagged dependent variable and instruments for central government expenditure using log population and distance from district capital to Lusaka. As five education outcomes are time invariant, we were able to estimate OLS, 2SLS, GMM, and LIML models on 18 welfare outcomes whereas only on 13 outcomes using SGMM models.

also the step-down approaches of Holm (1979) and Holland and Copenhaver (1987).

A drawback of multi-test procedures that control for FWER is that they are, as pointed out by Newson and The Alspac Study Team (2003), very conservative and can result in low power to detect real differences. Therefore, we also resort to less conservative procedures to control for the false discovery rate (FDR), including the step-down methods of Benjamini and Liu (1999a, 1999b), and the step-up approach of Benjamini and Yekutieli (2001). Given that the *p*-values of our sample are not normally distributed, and that they could, in principle, be negatively correlated, we focus on the results from the Holm and Benjamini and Yekutieli methods, which are more conservative procedures that allow us to control for FWER and FDR, respectively.[22]

We calculate these procedures by setting an upper bound to the set of rejected null hypotheses that are true for FWER and the FDR at 0.05 levels. The results are presented in Table 7. They indicate that for each of the rejected hypotheses we can be 95% confident that they are not a false discovery, i.e., do not represent a type I error.

6. DISCUSSION

Here we consider several possible explanations for the Zambian results, and in particular the result most directly counter to what the diversity debit hypothesis predicts – the statistically significant *positive* relationship found between ethnic diversity and welfare outcomes.

This positive association could be consistent either with diversity driving improved outcomes, or improved outcomes driving higher diversity. Only the first implies a diversity dividend, while the second could be consistent with diversity debit which is masked by the other, larger effect.

Arguably the strongest argument against a diversity dividend relates to the possible endogeneity of ethnic diversity to various outcomes. Interestingly, the argument that ethnic diversity is endogenous has also received a lot of attention in critiques of quantitative studies of a diversity debit for growth and development. Although much of this work has assumed that ethnic diversity is exogenous, it is well-established in the literature on ethnic politics that ethnic groups and divisions are constructed and influenced precisely through the processes explored in such analyses (Chandra, 2012; Green, 2013). Modernization and development, for instance, may influence the emergence of broader national identities and the weakening of "traditional" ethnic affiliations, consistent with a correlation between high diversity in poor countries and low diversity levels in wealthier countries.

Within the context and timeframe explored in our analysis, however, a more likely way in which diversity would be endogenous has to do with migration – i.e., if better public goods provision in certain districts attracts immigration from worse performing districts. Gerring *et al.* (2015) likewise highlight this possibility as one explanation for their results, positing that such "optimal sorting" is more likely at the sub-national level because people can more freely move within a country than across national borders, making diversity at sub-national levels more a matter of choice than birth. At the sub-national level "people who wish to live together are more likely to be able to do so" and "those who choose to relocate to diverse areas – rather than staying put or relocating to an ethnic enclave comprised of persons with similar backgrounds – are likely to be more ambitious, more skilled, and more highly educated (Borjas, 1998; Damm, 2009) – and, one might add, less averse to living amidst diversity" (Gerring *et al.*, 2015, p. 174).

While some optimal sorting may be occurring in Zambia, we doubt that this is what is driving our results. For one, our analysis excludes the four cities in the sample – presumably where the most ambitious, pro-diversity migrants would prefer to go. To the extent that we expect those in rural and municipal districts not to be so different from each other, optimal sorting seems unlikely to be a major explanatory factor. Further, additional analysis of our data does not support the hypothesis of optimal sorting as differences in the quality of public services across districts do not seem to be a relevant factor in internal migration patterns, and migration does not have a substantial impact on ethnic diversity at district level.

Both the LCMS and the population census record whether and why individuals migrated into a district during the 12 months prior to the survey. The 2006 LCMS shows that internal migration in Zambia is fairly low, with about 3% of individuals (and households) having relocated during the past 12 months, the vast majority of which relocated within the same district – i.e., their mobility would not have changed ethnic fractionalization measures at district level (Zambia, 2006, p. 27). More importantly, differentials in the quality of public services between districts do not appear to be a major driver of rural inter-district migration. For instance, only 6% in the relevant age group of 12–19 years gave attending school as a reason for relocating (2% in the age group 0–11).

Analysis of census data does not support the hypothesis of optimal sorting either. Using information from the 2000 and the 2010 population census, we were able to calculate the correlation between public goods provision and the observed changes in ethnic fractionalization for all and "major" ethnic groups in the period 2000–10. Our results show very low correlations for health and education outcomes, the highest for primary school enrollment at $r = 0.1051$ for major groups, and for FIC immunization at $r = 0.0828$ for all groups. It thus seems fair to say that our results are not driven by reverse causation running from health and education outcomes to ethnic fractionalization.

A related argument, put forward by Bates (1974) in the context of African urbanization, but arguably applicable to the rural context as well, is that processes of modernization themselves have served to increase ethnically-based identification and competition over education and jobs in particular. This suggests that populations in more heterogeneous communities may demand and make use of public services more actively than in more homogeneous ones, for instance as a result of more intense inter-group competition in the education system and the local labor markets.[23]

Another possible explanation for our results is how 'scale' may play a role in making sub-national communities more likely to realize the benefits of diversity (Gerring *et al.*, 2015). For example, we see in other contexts that local communities can coordinate to provide common pool resources, but such informal mechanisms may be unlikely to function at the national level (Ostrom, 1990). With respect to Zambia, however, the size of the sub-national units in question (districts) raises questions about whether community politics at this scale really benefit from the sort of coordination Ostrom shows: her focus is explicitly on 'small-scale CPRs [common pool resources], where . . . the number of individuals affected varies from 50 to 15,000 persons who are heavily dependent on the CPR for economic reasons' (Ostrom, 1990, p. 26). District-level governance in Zambia operates on a larger scale and without the same economic incentives.

Gibson and Hoffman (2013, p. 273) offer yet another explanation that "political institutions" – namely, electoral systems – "can create incentives for politicians to work across ethnic lines, even where ethnicity is a salient political factor." To be re-elected, politicians need to deliver benefits to constituents, which in turn provide incentives for them to form coalitions to pass policy. Building on Bawn and Rosenbluth (2006)'s findings on political fragmentation and government expenditure, they propose that just as public expenditure increases with the number of parties, it should also increase with ethnic diversity. While this argument offers important traction on explaining the district-level spending analyzed by Gibson and Hoffman, it does not offer clear predictions with respect to our finding that central government expenditure is negatively correlated with ethnic fractionalization, whereas outcomes improve with fractionalization.

Interestingly, however, Gibson and Hoffman's empirical findings suggest another explanation for our finding of a statistically significant *negative* relationship between ethnic fractionalization and budget allocations: As noted above, we would not have expected district-level ethnic fractionalization to *directly* influence national budget allocations. However, it might do so indirectly. In particular, to the extent that ethnic fractionalization has an observed positive relationship with district-level outputs and outcomes, it may indirectly lead to lower allocations to these "less needy" districts. Our analysis is supportive of this argument as we find central government expenditure to be endogenous in our econometric model.[24]

Another interpretation of our findings links to the extensive literature on neopatrimonial regimes in Africa. The concept of neopatrimonialism stresses the lack of effective checks and balances in the public sector, the importance of informal rules and institutions for the distribution of public resources and the capture of these resources by elites and leaders to maintain extended clientelistic networks and patronage systems (Bratton & van de Walle, 1994; Erdmann & Simutanyi, 2003; Leiderer *et al.*, 2007; von Soest, 2007). If, in the absence of effective accountability mechanisms in the formal governance system, resources transferred from central government in Zambia are subject to such capture by local groups, then one might expect informal local rules and institutions to be important determinants of the extent to which such capture takes place in a particular district. Various studies suggest that some form of local capture through informal processes may indeed be happening at district level in Zambia. In a World Bank Public Expenditure Review for the Zambian health sector in 2008, Picazo and Zhao (2009, p. 33) argue that while it is possible to trace resources from the Ministry of Health to the districts, how allocation decisions are taken within the district health management teams remains a "black box" and does not form part of the formal "fiscal information chain." In a limited sample their study finds that only 50% of reviewed health centers received their full allocations (see Leiderer *et al.*, 2012, p. 131).

Several informal processes would be consistent with the positive relationship we find between ethnic diversity and public goods outcomes. For instance, if one assumes that central government transfers captured by local leaders are not used for the provision of public goods but mainly for private consumption and patronage spending on each leader's own group, then each group might have a strong incentive to curtail such capture by competing groups. It can be expected that whether one group can effectively keep another other one from misappropriating public resources aimed at funding public goods, strongly depends on the relative size of these groups.

One would thus expect "informal" checks and balances at local level to work more effectively across more diverse communities of comparable size than when a local community is dominated by only one large group or a small number of groups that may collude in diverting public resources. In this case one would expect capture of public resources for private consumption or patronage spending to be more prevalent in more homogeneous districts than in those with a more diverse population. *Vice versa*, in more heterogeneous communities, where informal checks and balances between ethnic groups and their traditional leaders exist, and no single dominant group (or their leader) is able to capture a major share of central government transfers, a larger share of central government transfers will be available for the provision of public goods and services. As a result, one would thus expect better welfare outcomes for given levels of central government spending in ethnically more diverse communities than in more homogeneous ones. This line of argument seems to be consistent with the recent experimental findings of Levine *et al.* (2014) which suggest that ethnic diversity prevents detrimental herding behavior and fosters greater inter-group scrutiny, which in turn leads to better outcomes.

In sum, none of these diverse explanations alone is fully satisfactory in explaining our empirical findings, but together they speak to why ethnic diversity does not necessarily undermine public goods provision and the possible channels that may underlie a diversity dividend.

7. CONCLUSION

The findings in this study challenge the conventional wisdom that ethnic diversity is associated with the under-provision of public goods, a hypothesis that has been applied particularly to understanding developmental deficits in sub-Saharan Africa. Using district-level data for Zambia across a wider range of public goods-related indicators than analyzed in previous work on the region, we show that ethnic fractionalization is not associated with the under-provision of public goods and, indeed, has a *positive* relationship with some key welfare outcomes. Our results are consistent across alternative specifications and models. Drawing on these results and also an emerging body of work challenging the diversity debit hypothesis on empirical grounds, we conclude, contra the conventional wisdom, that given the more nuanced empirical relationships now documented here, the key question for future work is not so much "why does ethnic diversity undermine public goods provision," but "when does ethnic diversity support public goods provision and aggregate welfare and when does it undermine it?" and "why and under what conditions does a diversity divided exist?" (Habyarimana *et al*., 2007). What are the processes or mechanisms underlying the diversity dividend that has now been documented in multiple sub-national empirical studies, including our own?

We do not have good answers to these questions, yet. However, by critically considering several possible explanations proposed in the literature against the Zambian data, we point to some promising directions for future research.

NOTES

1. Ethnic diversity may also be endogenous to development outcomes. We return to this point later in the discussion.

2. It is worth nothing that these arguments are not necessarily tied to the ethno-cultural characteristics of ethnic groups. If ethnic groups tend to be regionally concentrated – for whatever reason – they may have different and conflicting interests over where to locate schools, roads, or health centers, regardless of cultural commonalities, ethnic hatreds, or historical ethnic myths of origin. Likewise, economic inequalities between groups may also drive the relationship, perhaps through impact on between-group differences in preferences, prejudice, and social capital (Alesina & La Ferrara, 2000; Baldwin & Huber, 2010; Waring, 2012).

3. Private consumption (c) is equal to exogenous pre-tax income (y) minus a lump-sum tax (t). The population size is normalized at 1, so that g represents the per capita and aggregate size of the public good.

4. Why this relationship holds however remains open for discussion. Alesina *et al*.'s (1999) model is widely cited, but it relies on the median voter theorem which is generally formulated on the basis of two-party competition under plurality rule – i.e., a different institutional context to many of the countries under analysis. Interpretation is further complicated by the quality of data on public goods provision and government budgets that is available at the cross-national level.

5. In the rest of this paper, we use "diversity" and "divisions" interchangably. In referring to diversity, we mean more specifically diversity measured using socially and politically salient categories such as those used in our Zambian analysis.

6. There is also a positive relationship with spending on police. This arguably can be reconciled within the model: polarized preferences may also lead to higher levels of social conflict and thus greater demands for policing.

7. For instance, intergovernmental transfers in the US may be influenced indirectly by ethnic diversity, thus affecting local spending.

8. Nor, in most of Africa, do they contribute a large share of tax revenues.

9. Although we have data on welfare outcomes for the period 2001–09, the shorter time window of the expenditure data, from 2004 to 2009, meant that we were able to only cover the period 2004–09 in our analysis.

10. It is worth noting that the ethnic groups used in our analysis also approximate the politically salient groups identified by the Ethnic Power Relations dataset version 3.01 (Wimmer, Cederman, & Min, 2009): Bemba speakers, Tonga-Ila-Lenje, Nyanja speakers (Easterners), Lozi (Barotse), Lunda (NW Province), Luvale (NW Province), and Kaonde.

11. In spite of their similarity, it can be argued that each index reflects different aspects of ethno-linguistic diversity. The ELF-E is presumably the more accurate measure of fractionalization in terms of ethnic self-identification and thus speaks to those mechanisms underlying the diversity debit hypothesis that stress the role of preferences or trust within and across ethnic groups. The ELF-L in contrast may be more relevant to the social salience of fractionalization in terms of language use for "day-to-day communication with [...] neighbors, at factory, in office, in market places, etc." (Republic of Zambia, 2000, p. 43), which is arguably driven at least partly by social and economic needs rather than ethnic self-identification. This measure therefore may be more relevant to those mechanisms that stress the role of social capital. As is to be expected, the average ELF-L (and its standard deviation) is substantially smaller than the ELF-E (see Table 1), as people in ethnically diverse communities will generally use only a limited number of common languages for everyday communication.

12. These grants are used by basic schools to purchase mostly locally procured learning and teaching materials.

13. Population by age-group is only available from the 2000 Census.

14. We note that SGMM may suffer from the weak instrument problem when the time series is large and substantial unobserved heterogeneity exists (Bun & Windmeijer, 2010; Hayakawa, 2007). Given the short time series of our data, we suspect this problem to be minimal. Another potential deficiency of the SGMM estimators is that the number of internal instruments grows quadratically as the number of time periods increases. Roodman (2009) cautions that instrument proliferation can over-fit endogenous variables, biasing coefficient estimates and weakening the Hansen test of the instruments' joint validity. Therefore, we reduce the instrument count by "collapsing" instruments which is superior to simply restricting the lag ranges.

15. Detailed regression results for ELF-E are reported in Appendix C.

16. See Tables C4, C6, C7 in Appendix B.

17. See Tables C2, C4, C7 in Appendix C.

18. Results not reported but available on request.

19. Details not reported but available on request.

20. Budget execution rates vary substantially between districts and years. For overall education expenditure, the average execution rate in the sample is 99.3%, with a standard deviation of 8.5 percentage points and a minimum value of 51.6% and a maximum of 126%. For health expenditure, the mean execution rate is 66% (standard deviation 29 percentage points), with a minimum value of 1% and a maximum of 130%.

21. Including the ruling party's vote share as an additional endogenous regressor does not substantially alter the results, but slightly increases coefficients on ethnic fractionalization and significance levels. In addition, the Sargan–Hansen tests of overidentifying restrictions perform somewhat better in individual specifications. The results reported here, therefore, arguably represent conservative estimates of the effect of ethnic fractionalization.

22. Note that other procedures produce very similar results.

23. This argument appears broadly consistent with our finding that in particular outcomes that require the active participation of target groups, such as primary school enrollment and children's immunization rates, are positively correlated with ethnic fractionalization.

24. As Gibson and Hoffman's (2013) analysis suggests, there may also be a positive relationship between ethnic fractionalization and district council revenues and spending, which – even if only a minor share of the total budget – could also help to explain the negative sign in our results on allocations.

REFERENCES

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics, 114*, 1243–1284.

Alesina, A., & La Ferrara, E. (2000). Participation in heterogeneous communities. *The Quarterly Journal of Economics, 115*, 847–904.

Anderson, D., & Cheeseman, N. (2013). *Routledge handbook of African politics*. Routledge.

Ashraf, Q., & Galor, O. (2013a). Genetic diversity and the origins of cultural fragmentation. *American Economic Review, 103*, 528–533.

Ashraf, Q., & Galor, O. (2013b). The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development. *American Economic Review, 103*, 1–46.

Bahry, D., Kosolapov, M., Kozyreva, P., & Wilson, R. K. (2005). Ethnicity and trust: Evidence from Russia. *The American Political Science Review, 99*, 521–532.

Baldwin, K., & Huber, J. D. (2010). Economic versus cultural differences: Forms of ethnic diversity and public goods provision. *American Political Science Review, 104*, 644–662.

Banerjee, A., Iyer, L., & Somanathan, R. (2005). History, social divisions, and public goods in rural India. *Journal of the European Economic Association, 3*, 639–647.

Bates, R. (1974). Ethnic competition and modernization in contemporary Africa. *Comparative Political Studies*, 457–483.

Bawn, K., & Rosenbluth, F. (2006). Short versus long coalitions: Electoral accountability and the size of the public sector. *American Journal of Political Science, 50*, 251–265.

Benjamini, Y., & Liu, W. (1999a). *A distribution-free multiple test procedure that controls the false discovery rate*. Tel Aviv University.

Benjamini, Y., & Liu, W. (1999b). A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence. *Journal of Statistical Planning and Inference, 82*, 163–170.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics, 29*, 1165–1188.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*, 115–143.

Borjas, G. J. (1998). To ghetto or not to ghetto: Ethnicity and residential segregation. *Journal of Urban Economics, 44*, 228–253.

Bossert, T., Chitah, M. B., & Bowser, D. (2003). Decentralization in Zambia: Resource allocation and district performance. *Health Policy and Planning, 18*, 357–369.

Bratton, M., & van de Walle, N. (1994). Neopatrimonial regimes and political transitions in Africa. *World Politics, 46*, 453–489.

Chandra, K. (2001). Ethnic bargains, group instability, and social choice theory. *Politics and Society, 29*, 337–362.

Chandra, K. (2012). *Constructivist theories of ethnic politics*. New York: Oxford University Press.

Cox, G. W., & McCubbins, M. D. (1986). Electoral politics as a redistributive game. *The Journal of Politics, 48*, 370–389.

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics, 27*, 281–314.

Davidson, J. (2000). *Econometric theory*. Oxford: Wiley-Blackwell.

De Kemp, A., Faust, J., & Leiderer, S. (2011). *Between high expectations and reality: An Evaluation of budget support in Zambia (2005–2010)* (Vol. Synthesis report 356). Bonn/The Hague/Stockholm: German Federal Ministry for Economic Cooperation and Development/Ministry of Foreign Affairs of The Netherlands/Swedish International Development Cooperation Agency.

Deutsch, K. (1966). *Nationalism and social communication*. Cambridge: MIT Press.

Dixit, A., & Londregan, J. (1996). The determinants of success of special interests in redistributive politics. *The Journal of Politics, 58*, 1132–1155.

Driscoll, J. C., & Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics, 80*, 549–560.

Easterly, W., & Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *The Quarterly Journal of Economics, 112*, 1203–1250.

Easterly, W., & Rebelo, S. (1993). Fiscal policy and economic growth. *Journal of Monetary Economics, 32*, 417–458.

Erdmann, G., & Simutanyi, N. (2003). *Transition in Zambia. The hybridisation of the third republic*. Lilongwe: Konrad Adenauer Foundation.

Franck, R., & Rainer, I. (2012). Does the leader's ethnicity matter? Ethnic favoritism, education, and health in Sub-Saharan Africa. *American Political Science Review, 106*, 294–325.

Gebregziabher, F., & Niño-Zarazúa, M. (2014). *Social spending and aggregate welfare in developing and transition economies*. WIDER working paper series, 2014/082.

Gerring, J., Thacker, S. C., Lu, Y., & Huang, W. (2015). Does diversity impair human development? A multi-level test of the diversity debit hypothesis. *World Development, 66*, 166–188.

Gibson, C. C., & Hoffman, B. D. (2013). Coalitions not conflicts: Ethnicity, political institutions, and expenditure in Africa. *Comparative Politics, 45*, 273–290.

Gisselquist, R. M. (2014). Ethnic divisions and public goods provision, revisited. *Ethnic and Racial Studies, 37*, 1605–1627.

Golden, M., & Min, B. (2013). Distributive politics around the world. *Annual Review of Political Science, 16*, 73–99.

Gören, E. (2014). How ethnic diversity affects economic growth. *World Development, 59*, 275–297.

Gormley, W. T. (2007). Public policy analysis: Ideas and impacts. *Annual Review of Political Science, 10*, 297–313.

Green, E. (2013). Explaining African ethnic diversity. *International Political Science Review, 34*, 235–253.

Habyarimana, J., Humphreys, M., Posner, D. N., & Weinstein, J. M. (2007). Why does ethnic diversity undermine public goods provision?. *American Political Science Review, 101*, 709–725.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46*, 1251–1271.

Holland, B. S., & Copenhaver, M. D. (1987). An improved sequentially rejective Bonferroni test procedure. *Biometrics*, 417–423.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

ILO (2008). *Zambia: Social protection expenditure and performance review and social budget*. Geneva: International Labour Organization.

IOB (2008). *Primary education in Zambia* (Vol. 312). The Hague: Policy and Operations Evaluation Department.

Jackson, K. (2013). Diversity and the distribution of public goods in sub-saharan Africa. *Journal of African Economies, 22*, 437–462.

Jerven, M. (2013). *Poor numbers: How we are misled by African development statistics and what to do about it*. Ithaca: Cornell University Press.

Kanbur, R., Rajaram, P. K., & Varshney, A. (2011). Ethnic diversity and ethnic strife. An interdisciplinary perspective. *World Development, 39*, 147–158.

Khwaja, A. I. (2009). Can good projects succeed in bad communities?. *Journal of Public Economics, 93*, 899–916.

Kimenyi, M. S. (2006). Ethnicity, governance and the provision of public goods. *Journal of African Economies, 15*, 62–99.

Kramon, E., & Posner, D. N. (2013). Who benefits from distributive politics? How the outcome one studies affects the answer one gets. *Perspectives on Politics, 11*, 461–474.

Leiderer, S. (2014). *Who gets the schools? Political targeting of economic and social infrastructure provision in Zambia* (Vol. Discussion paper 27/2014). Bonn: German Development Institute/Deutsches Institut für Entwicklungspolitik (DIE).

Leiderer, S., Geigenmüller, M., Hornig, A., Kästle, K., Smith, C., & Tröger, F. (2012). *Efficiency of service provision at local government level in Zambia in health, education and roads: Implications for decentralisation and the effectiveness of budget support* (Vol. Studies 71). Bonn: German Development Institute/Deutsches Institut für Entwicklungspolitik (DIE).

Leiderer, S., Hodick, B., Kabey, E., Roll, M., Schnitzer, S., & Ziegenbein, J. (2007). *Public financial management for PRSP implementation in Malawi – Formal and informal PFM institutions in a decentralising system* (Vol. Studies 28). Bonn: German Development Institute/Deutsches Institut für Entwicklungspolitik (DIE).

Levine, S. S., Apfelbaum, E. P., Bernard, M., Bartelt, V. L., Zajac, E. J., & Stark, D. (2014). Ethnic diversity deflates price bubbles. *Proceedings of the National Academy of Sciences*.

Lieberman, E. S., & McClendon, G. H. (2013). The ethnicity-policy preference link in Sub-Saharan Africa. *Comparative Political Studies, 46*, 574–602.

Lijphart, A. (2012). *Patterns of democracy: Government forms and performance in thirty-six countries*. New Haven: Yale University Press.

Lindbeck, A., & Weibull, J. W. (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice, 52*, 273–297.

Lindemann, S. (2011a). The ethnic politics of coup avoidance: Evidence from Zambia and Uganda. *Africa Spectrum, 46*, 3–41.

Lindemann, S. (2011b). Inclusive elite bargains and the dilemma of unproductive peace: A Zambian case study. *Third World Quarterly, 32*, 1843–1869.

Mason, N. M., Jayne, T. S., & van de Walle, N. (2013). *Fertilizer subsidies and voting patterns: Political economy dimensions of input subsidy programs*. Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4–6, 2013.

Mavridis, D. (2015). Ethnic diversity and social capital in Indonesia. *World Development, 67*, 376–395.

McDoom, O. S., & Gisselquist, R. M. (2015). The measurement of ethnic and religious divisions: Spatial, temporal, and categorical dimensions with evidence from mindanao, the Philippines. *Social Indicators Research*. http://dx.doi.org/10.1007/s11205-015-1145-9.

Miguel, E. (2004). Tribe or nation? Nation building and public goods in Kenya versus Tanzania. *World Politics, 56*, 327–362.

Miguel, E., & Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in Kenya. *Journal of Public Economics, 89*, 2325–2368.

Newson, R., & The Alspac Study Team (2003). Multiple-test procedures and smile plots. *Stata Journal, 3*, 109–132.

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge: Harvard University Press.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge: Cambridge University Press.

Picazo, O., & Zhao, F. (2009). *Zambia health sector public expenditure review*. Washington, DC: World Bank.

Posner, D. N. (2004). Measuring ethnic fractionalization in Africa. *American Journal of Political Science, 48*, 849–863.

Posner, D. N. (2005). *Institutions and ethnic politics in Africa*. Cambridge University Press.

Putnam, R. D. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton: Princeton University Press.

Putnam, R. D. (2007). E Pluribus Unum: Diversity and community in the twenty-first century: The 2006 Johan Skytte prize lecture. *Scandinavian Political Studies, 30*, 137–174.

Reinikka, R., & Svensson, J. (2004). Local capture: Evidence from a central government transfer program in Uganda. *The Quarterly Journal of Economics, 119*, 679–708.

Republic of Zambia (2000). *Zambia 2000 census of population and housing: Enumerator's instruction manual*. Lusaka: C. S. Office.

Republic of Zambia. (2008). Zambia public financial management report and performance indicators. 2008 Assessment and update.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics, 71*, 135–158.

Singh, P. (2010). We-ness and welfare: A longitudinal analysis of social development in Kerala, India. *World Development, 39*, 282–293.

Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In W. K. A. Donald (Ed.), *Identification and inference for econometric models* (pp. 80–108). New York: Cambridge University Press.

van de Walle, N. (2001). *African economies and the politics of permanent crisis, 1979–1999*. New York: Cambridge University Press.

von Soest, C. (2007). How does neopatrimonialism affect the African state's revenues? The case of tax collection in Zambia. *The Journal of Modern African Studies, 45*, 621–645.

Wantchekon, L. (2003). Clientelism and voting behavior: Evidence from a field experiment in Benin. *World Politics, 55*, 399–422.

Waring, T. M. (2012). Cooperation dynamics in a multiethnic society: A case study from Tamil Nadu. *Current Anthropology, 53*, 642–649.

Wimmer, A., Cederman, L.-E., & Min, B. (2009). Ethnic politics and armed conflict: A configurational analysis of a new global data set. *American Sociological Review, 74*, 316–337.

World Bank. (2008). Making basic education services more effective in Zambia: Improving delivery, tracking school level expenditures. Washington, DC.

Zambia, R. O. (2006). *Living condition monitoring survey report 2006*. Lusaka: Central Statistical Office.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.worlddev.2015.10.018.

Available online at www.sciencedirect.com

ScienceDirect