

Summary

This paper uses the methodology of Fajgelbaum and Schaal (2017) (FS17 hereafter) to compute the optimal transport network for each african country taken individually in a spatial trade model that features 2 goods: an urban and an agricultural good. The model is estimated on data coming from Open Street Map (current road network), night lights from Henderson et al. (2012) (GDP/productivity) and GPW (population). The optimal network is then computed for each country and the authors constructs a measure of infrastructure misallocation by evaluating the welfare gains of moving from the current network to the optimal one, as well as a local measure of *infrastructure discrimination* which is the ratio of the welfare in the optimal vs. current allocation in a single location.

In the second part of the paper, the authors seeks to understand the determinants of infrastructure misallocation by regressing the local discrimination measure to i) infrastructure investments by previous colonial power, ii) measures of potential ethnic discrimination and iii) foreign aid. The results suggest an important role for colonial investments and an almost negligible role for ethnic discrimination. Foreign aid, on the other hand, seems inefficiently directed to regions that are already overly privileged, according to the model.

General assessment

There is a lot to like about this paper and I really enjoyed reading it. An impressive amount of work went into bringing together the different datasets and implementing the optimal network calculations. This being said, the first part of the paper mostly remains an application of an existing methodology and I am not sure the contribution is sufficiently novel for a top journal like the AEJMacro. The second part, on the other hand, is more novel in this kind of quantitative framework but the identification effort remains a little light and one is left to wonder in the end what the model truly brings for that part (many of these regressions had been conducted in the literature before using direct measures of infrastructure investments: what exactly are we gaining with this new discrimination index?). The final results also point towards various limitations of the FS17 framework and I would like to invite the author to be more critical in its application and consider additional robustness tests.

Comments

1. *Modeling choices.* I have some concerns about several modeling choices you made in applying the FS17 methodology.
 - (a) *Two goods.* While I understand the need to restrict the number of goods to reduce the number of variables over which you run the optimization in the optimal network calculations, the assumption that there exists only two goods

is extreme and perhaps too particular. One particular implication is, most likely, that cities do not want to trade with each other. I do not know enough the nature of interregional trade in Africa, but we know for other regions that within-industry trade in differentiated product can be quite high. Your assumption is, I would guess, not innocuous for the shape of the optimal network: isn't it true in that case that trade takes place mostly between a city and its surrounding hinterland and that the planner would simply try to partition the land between different autonomous regions that would have no trading relationships? I would strongly encourage the author to explore other ways to specify the production and good allocation to evaluate the robustness of the model's current predictions. While the Armington assumption might be too extreme (and impossible to implement), one way could be to follow FS17 in assuming one distinct differentiated good for the largest city and maintain the homogeneous good assumption for other locations.

- (b) *No mobility.* You assume that population is immobile. While this may be acceptable for distant and isolated regions, there is also evidence of large migrations between cities and rural areas. Since the FS17 framework allows for mobility, could you check how different your discrimination measure would be to relaxing that assumption?
 - (c) *No international trade.* Another important restriction is that you compute the optimal transport network for each country individually ignoring how much a country trades with each neighbor. This may result in substantial misspecification especially for regions along national borders. Could you compute the optimal network without making this overly restrictive assumption? (you could perhaps take as given some demand functions from other countries at the main ports of entry? or compute the overall optimal transport network for Africa as a whole?)
 - (d) *Convex vs. nonconvex.* The elasticity γ of trade costs with respect to infrastructure investments is a little arbitrary. You choose a value that guarantees the convexity of your planning problem but without any empirical support. Gilles Duranton and coauthors have recent work trying to estimate this elasticity.
 - (e) *Other modes of transportation.* In conducting the estimation, you ignore other modes of transportation like water transport and freight by train. While including those modes of transportation in the model is probably asking for too much, could you tell us at least how much trade is currently conducted by roads vs. other modes?
2. *Beyond Fajgelbaum and Schaal (2017).* The FS17 methodology allows you to compute the globally optimal network, but this comes at the cost of making a number of restrictive assumptions. While it may be beyond the scope of that paper to relax these assumptions, I would still like to invite the author to think

about them and, to the very least, discuss how much bias they may introduce into the results.

- (a) *Exogeneity of productivity and absence of agglomeration forces.* Your whole discussion of the effect colonial investments and foreign aids points towards one key limitation of the model which is that productivity is exogenous. Economic geography has shown over the years the importance of increasing returns and agglomeration forces. While writing a version of the model with dynamics and endogenous growth is probably beyond reach at the moment, could you find a way to tell us how your results would change in the presence of simple agglomeration forces (for instance, local spillovers in productivity). The framework of Allen and Arkolakis (2019), while only local, seems to be able to handle some externalities of the kind. Could you combine the two approaches and evaluate the optimal network or welfare gradient in that case?
 - (b) *Efficiency of allocation.* The FS17 approach computes the planner's allocation not only for the infrastructure investment but also for the entire economic allocation. This assumes, in particular, that the planner has access to a full set of Pigovian taxes to offset the congestion externality as well as transfers between regions. This further assumes that there can be no distortions in the product market (market power, etc.). You may or may not be able to do much in that case, but could you tell us in any way how your results would change if the allocation was not efficient to start with?
3. *Empirical part.* I found the results from the empirical part in section 5 quite fascinating. But I think there is some room for improvement.
- (a) *Role of the model.* As I mentioned at the beginning of this report, there is a lingering feeling in the empirical section that one doesn't really know what the model is used for. Another alternative could have been to run the exact same regressions on existing measures of the infrastructure stock or investment (perhaps including the model's predicted investment as one of the regressor). What the model does, if we believe it, is that it controls *correctly* for the way fundamentals should predict infrastructure investments. Indeed, capturing how fundamentals across locations interact in a network to predict infrastructure investments cannot be simply controlled by controls like vector $X_{i,c}$. So the model should tell us how we should do that for productivity, population, network centrality, etc. But then, one is left to wonder how we should take into account the forces that the model left out. The author himself seems to be confused as to whether the controls $X_{i,c}$ should include measures like population, output, ruggedness that are already included into the model (page 26). So overall, the objective of the exercise is a little unclear and more thought should be given to what we want to do with the model and how we should treat model misspecification. The literature on misallocation may give a potential answer to that.

- (b) *Interpretation.* The interpretation of your results is sometimes a little unclear. Should we be surprised that colonial investments are negatively associated with the discrimination index? Isn't this effect quite mechanical? If regions with high colonial investments start with a higher stock of infrastructure, should we be surprised that the planner would want to redistribute those resources to other regions? There is also the fact that colonial investments may be correlated to economic fundamentals that are still relevant today. I think it would be more interesting to look at colonial investments that are orthogonal to what the optimal allocation would require and see to what extent it still contributes to infrastructure misallocation today. Perhaps running the regression at the national level could be interesting as well (how much total colonial investment resulted in misallocation as a whole). I was also confused about the foreign aid regressions. Does the dataset you are using merely reflect currently ongoing projects? Could you say something about the total aid received by a region over the past decades? I think it would be interesting to split your estimates between entirely new projects that have not yet started with other projects that have been ongoing for a long time (in which case a negative estimate could simply reflect the resulting higher stock of infrastructure).
- (c) *Identification.* Your OLS regressions tend to merely reflect correlation between the discrimination measure and the variables at hand. It would nice to think more seriously about causation. For instance, did foreign aid lead to a very unequal/distorted distribution of infrastructure? Or did an unequal distribution of power across region attract foreign aid in a very distorted way? Similarly, colonial investments may have been themselves a function of fundamentals, so could you come up with an instrument that would control for that?