



Definition matters. Metropolitan areas and agglomeration economies in a large-developing country[☆]

Maarten Bosker^a, Jane Park^b, Mark Roberts^{c,*}

^a Department of Economics, Erasmus University Rotterdam, The Netherlands and CEPR

^b Urban, Disaster Risk Management, Resilience and Land Global Practice, The World Bank, Washington, DC, USA

^c Urban, Disaster Risk Management, Resilience and Land Global Practice, The World Bank, Singapore



ARTICLE INFO

JEL Codes:

018
047
C21

Keywords:

Metro areas
Urban definitions
Agglomeration economies
Indonesia

ABSTRACT

A variety of approaches to delineating metropolitan areas have been developed. Systematic comparisons of these approaches in terms of the metro area landscape that they generate are however few. Our paper aims to fill this gap. We focus on Indonesia and make use of data on commuting flows, spatially fine-grained population, and remotely sensed nighttime lights to construct metropolitan areas using several approaches that have been developed in the literature. We find that the maps and characteristics of Indonesia's metro area landscape generated when using a commuting flow approach differ substantially from those generated using other approaches. Moreover, combining information on the metro areas generated by the different approaches with detailed micro-data from Indonesia's national labor force survey, we show that the estimated agglomeration wage premium for Java-Bali tends to fall when using a more restrictive definition of metro areas. This is not the case for the rest of Indonesia, for which we, moreover, find a much lower estimated agglomeration wage premium. We provide an explanation for these findings, and also tentatively probe the factors behind Indonesia's agglomeration wage premium.

1. Introduction

Traditionally, urban economists have relied on data for cities as defined by their administrative boundaries or, where available, by their official metropolitan boundaries. However, administrative boundaries often fail to adequately delineate a city's extent. This is especially the case in situations where urbanization has been rapid and/or many cities have been sprawling in land area, leading them to spill-over, increasingly inaccurate, administrative boundaries (see, for example, [Ellis and Roberts, 2016](#)). Similarly, official metropolitan boundaries, while seemingly preferable to administrative boundaries, may be subject to biases owing to political pressures in their definition, aided by a lack of transparency in the algorithms used in their construction and the use of a multiplicity of *a priori* imposed criteria ([Duranton, 2015a](#)).

In reaction to this, there has been a recent growth in efforts to develop and apply algorithms that enable the better delineation of cities and metropolitan areas. These efforts have been led not just by economists (see, for example, [Rozenfeld et al., 2011](#); [Duranton, 2015a](#); [Baragwanath et al., 2020](#); [de Bellefon et al., 2020](#); [Dingel et al., 2020](#); [Henderson et al., 2020](#); [Galdo et al., 2020](#)), but also by both geographers and the remote sensing community, in which there is a long tradition of using satellite imagery to define a city's extent (see, for example, [Danko, 1992](#); [Elvidge et al., 1996](#)). Moreover, while academic economists have largely confined themselves to attempting to better delineate cities and metropolitan areas for individual countries or small groups of countries, international organizations such as the Development Bank of Latin America (CAF), the European Commission (EC), the Organization for Economic Co-operation and Development (OECD), and

* The authors thank Katie McWilliams, Benjamin Stewart, Andrii Berdnyk and Brian Blankespoor for their outstanding GIS support, as well as Jonathan Hasoloan and Shaun Zhang for supplemental GIS support. They also thank Massimiliano Cali and Giorgio Presidente for the sharing of data on manufacturing TFP. They further thank Gilles Duranton and two anonymous referees for their excellent comments. Participants at the 2019 European Urban Economics Association, the World Bank's 2019 Land and Poverty, and the 2019 Singapore Management University (SMU) Urban and Regional Economics conferences likewise provided very helpful comments, as did participants at seminars at the Dutch Bureau for Economic Policy Analysis and the World Bank. Financial support from both the Swiss State Secretariat for Economic Affairs (SECO) and the UK's Department for International Development (DFID) is also very gratefully acknowledged.

[☆] Corresponding author.

E-mail addresses: bosker@ese.eur.nl (M. Bosker), jpark16@worldbank.org (J. Park), mroberts1@worldbank.org (M. Roberts).

the World Bank have sought to develop and apply algorithms at regional and global scales (see, respectively, Ch et al., 2020; Dijkstra et al., 2019; Moreno-Monroy et al., 2020; World Bank, 2009). The ambition of these organizations has been to construct consistently defined regional and global data sets of cities or metropolitan areas to facilitate the uniform measurement of urbanization.

The preferred approach of economists to defining cities and metropolitan areas tends to be rooted in a labor market perspective based on commuting flow data, as with the definition of “statistical” metropolitan areas in the United States, France or Japan. This preference is consistent with the labor market pooling that occurs within cities being viewed (ever since Marshall, 1890) as one of the key sources of agglomeration economies. Also, the flow of workers is easier to track than either the flow of intermediate goods between firms or the flow of ideas (forward-backward linkages and knowledge spillovers being the other two prominent sources of agglomeration economies). Commuting patterns furthermore tend to occur over distances that people naturally recognize as metropolitan (Duranton, 2015a).

Commuting flow data, however, is hard to come by for many, especially developing, countries. Attempts to better delineate cities and metropolitan areas in such countries or to define regionally or globally consistent data sets of cities, therefore, instead rely on alternative approaches based on more readily available data. This includes the use of estimated travel times to approximate commuting sheds (Uchida and Nelson, 2009; World Bank, 2009; Ellis and Roberts, 2016; Moreno-Monroy et al., 2020) and approaches that associate cities with dense clusters of population (Dijkstra and Poelman, 2014; Dijkstra et al., 2019; Roberts, 2018a; Henderson et al., 2020). It also includes approaches that take advantage of satellite imagery to identify cities based on the amount of light they emit at night or their built-up area (Ellis and Roberts, 2016; Baragwanah et al., 2020; Dingel et al., 2020; Ch et al., 2020). Finally, there are approaches that apply algorithms to detailed maps, with precise locational information, of all buildings within a country to identify cities (Arribas-Bel et al., 2020; de Bellefon et al., 2020).

While, however, much effort has gone into developing and applying these, mainly morphological, approaches for the better delineation of cities and metro areas, only limited efforts have been made to compare the metropolitan area landscapes that they generate in terms of, for example, the number and sizes of metro areas identified. Also, little is known about how their use compares to an explicit commuting-based approach.¹ This lack of comparison is particularly interesting because, unlike a commuting based approach, the alternative approaches do not make use of any, or only limited, explicit information on actual flows of people, goods or ideas across space.

Given the above, the main aims of this paper are two-fold. First, we compare the results of different approaches for delineating metropolitan areas in terms of the basic descriptions they provide of the metro area landscape (notably, the number of metro areas, their size, the sub-national administrative units they consist of, and the shares of the urban population that they cover). Second, we explore whether, and if so how, key empirical results that are crucial to a proper understanding of the working of urban economies depend on the approach used to delineate these metropolitan areas. In particular, we show that the estimated agglomeration wage premium is sensitive to the choice of delineation approach.

More specifically, we compare four different approaches. The first is based on an algorithm developed by Duranton (2015a) and requires origin-destination (O-D) commuting flow data. The other three approaches are all non-commuting data approaches: the Agglomeration Index (AI), which was first developed by Uchida and Nelson (2009) for

the World Bank's 2009 World Development Report (World Bank, 2009); a “Cluster Algorithm” developed by Dijkstra and Poelman (2014) which associates cities with dense clusters of population and which forms the basis of the “degree of urbanization” approach that is discussed by Dijkstra et al. (2019); and, finally, the identification and delineation of metro areas based on the thresholding of *Night-Time Lights* data (NTL), similar to, for example, Ellis and Roberts (2016) and Dingel et al. (2020). The latter three approaches all have the advantage that they can be applied to any country in the world using readily available data sets, including those countries for which O-D commuting flow data is not available.

In all our empirics, we focus on the case of a single, large, developing country – Indonesia. We choose Indonesia for two main reasons. First, Indonesia's national labor force survey (*Survei Angkatan Kerja Nasional*, SAKERNAS) allows for the derivation of an O-D matrix of commuting flows between districts. As such, Indonesia is one of only a very few developing countries for which we can compare results obtained using a commuting data-based approach to the delineation of metropolitan areas to results from other approaches. Second, Indonesia, and especially its island-region of Java-Bali, provides a perfect example of a rapidly urbanizing developing country in which many cities have been expanding quickly in area, leading them to spill-over their administrative boundaries. This makes the choice of approach to delineating metropolitan areas that potentially straddle administrative boundaries particularly salient. In fact, there is a growing recognition among Indonesia's policymakers of the need for an appropriate algorithm to be able to identify “multi-district” cities or metropolitan areas (see, Roberts et al., 2019).

Of course, our focus on Indonesia raises the natural question of the extent to which our results may be taken as being of relevance to other countries. After all, Indonesia is an archipelago country consisting of more than 17,000 islands. Furthermore, a large share of its cities are coastal cities: 55% of its cities with a population of at least 300,000 lie within 15 km of a coastline.² However, while, *prima facie*, this may seem to bring into question the external validity of our exercise, it is important to remember that the bulk of Indonesia's urban, and, indeed, overall, population is concentrated on just four main island-regions. Java-Bali, Sumatra, Kalimantan and Sulawesi together accounted for 95.9% of Indonesia's overall urban population in 2016 with Java-Bali alone accounting for 68.9% (Roberts et al., 2019). Moreover, while it is true that a large share of its cities are coastal cities, Indonesia is by no means unique in this respect. For example, the share of cities with a population of 300,000 or more in the United Kingdom that are within 15 km of a coast (56%) is almost identical to the share for Indonesia. Meanwhile, the Philippines, Vietnam, Japan, Malaysia and Australia all have larger shares of their cities located within 15 km of a coast than does Indonesia.

In fact, from an external validity perspective, Indonesia's geography can be argued to be an advantage. In particular, its main island groupings exhibit very different population densities, each of which can be taken as representative of different parts of the developing world. A clear distinction can be drawn between Java-Bali and the rest of Indonesia in this regard. At 1093 people per km², Java-Bali had an average population density in 2016 close to that of Bangladesh, the Northeastern states of India, and the Nile delta. Meanwhile, outside of Java-Bali, no island-region had an average population density greater than 155 people per km² in 2016, which, according to data from the World Bank's World Development Indicators (WDI) database, is close to the average population density of all lower middle-income countries globally. In our analysis, we thus breakdown results between Java-Bali and the rest of Indonesia. This allows us to investigate the behavior and the com-

¹ Dingel et al. (2020) and Moreno-Monroy et al. (2020) in this special issue undertake comparisons of the specific approaches that they develop with a commuting-based approach.

² The calculations on the shares of cities that are within 15 km of a coastline were made using data from the United Nations' World Urbanization Prospects: 2018 Revision database (<https://population.un.org/wup/>).

parability of the different approaches in different population density settings.

By focusing on Indonesia, our paper also contributes to the emerging, but still relatively small, literature that provides credible empirical evidence on the strength of agglomeration economies in developing countries (Combes et al., 2015; Duranton, 2016; Chauvin et al., 2017; Combes et al., 2019; Quintero and Roberts, 2018; Roberts, 2018b; Henderson et al., 2020). This is a knowledge gap that Overman and Venables (2005), Duranton (2015b), and Glaeser and Henderson (2017) have all made calls for the profession to fill.

Finally, in examining how different approaches to metro area delineation affect the estimated agglomeration wage premium, our paper complements Briant et al. (2010) and Holmes and Lee (2010). Using French and US data respectively, these papers analyze how the size and shape of the spatial units used affects, *inter alia*, the measured degree of spatial concentration, the evidence on Zipf's Law, and the estimated agglomeration wage premium. What we add to these papers is a specific focus on the effects of the choice between different approaches to defining metropolitan areas, as well as of the choice of threshold(s) associated with any given approach, on the estimation of the agglomeration wage premium. Moreover, and consistent with the most recent empirical studies on the agglomeration wage premium (see Duranton, 2016; Chauvin et al., 2017; De La Roca and Puga, 2017), we analyze the effects of the choice of approach in a regression context where the outcome variable, a worker's nominal wage, is observed at the individual level. In this setting, differences in the delineation of metropolitan areas arising from the choice of approach lead to different allocations of workers across metropolitan areas, as well as between metro and non-metro urban areas. This, in turn, leads to differences in the measured level of density for these individual workers. It also leads to differences in the measurement of density for workers who are assigned to the same metro area by different approaches, but for which the approaches lead to the drawing of different boundaries for those metro areas.³

We find that the description of Indonesia's metropolitan area landscape produced can be very different depending on the approach and thresholds used to delineate metro areas. This is particularly so when comparing the metro areas defined by the commuting based approach with those defined by the other three approaches that we consider. On top of this, the estimated size of the agglomeration wage premium tends to fall when using a stricter approach and threshold(s) to delineate metro areas. This is, however, only the case on Java-Bali, where urbanization has been comparatively rapid with many fast-expanding cities spilling over administrative boundaries. In the rest of Indonesia, where urbanization is much less advanced and most urban areas remain well-confined within administrative districts, the agglomeration wage premium is almost invariant to the choice of approach and threshold used to delineate metro areas, as well as several orders of magnitude smaller than on Java-Bali. We provide explanations for these findings and, moreover, probe the possible factors behind Indonesia's agglomeration wage premium.

The remainder of the paper is structured as follows. Section 2 describes the four approaches for delineating metropolitan areas that we compare. Section 3 discusses the data that we use to apply these approaches to Indonesia. Section 4 provides a descriptive comparison of Indonesia's metropolitan area landscape under the different approaches. Section 5 presents our estimates of the agglomeration wage premium using the different approaches to defining metro areas. Section 6 provides

³ Importantly, we always keep the same fixed sample of (urban) workers (see Section 5 for more details). Workers that are not defined to be part of a metro, i.e. an urban area straddling administrative boundaries, are always retained in the sample. They are simply workers who are employed in urban areas whose boundaries are fully contained within a single administrative district.

a tentative empirical exploration of the sources of Indonesia's agglomeration economies. Section 7 concludes.

2. Approaches to defining metro areas

The four approaches to defining metro areas that we compare in this paper are:

Approach #1: Commuting-based approach – Duranton's commuting algorithm (2015a)

This approach identifies a metropolitan area as an integrated local labor market. All else equal in terms of data availability, such a functional approach to defining a city is frequently preferred by economists and is particularly appropriate for estimating the size and significance of the agglomeration wage premium based on labor force survey data, as we do in this paper. Duranton's algorithm (2015a) holds the advantage over other algorithms, such as those applied by the US Office of Management and Budget (2010) and OECD (2012), that similarly delineate metro areas based on their functional areas in that it does not require the pre-definition of metro cores nor the use of additional criteria beyond the specification of a simple commuting flow threshold (Duranton, 2015a). It is a simple iterative algorithm which uses sub-national administrative units (in our case, Indonesian districts) to "grow" metro areas through successive aggregation.

In the first round of running the algorithm, a district A is aggregated to a second district, B, if the share of workers that live in A and commute daily for work to B is above a given threshold, \bar{C} . In the second round, the algorithm then aggregates a third district, C, to the union of A and B, if the share of workers that live in C and commute daily to the spatial unit A + B exceeds \bar{C} , even though it may not have been possible to aggregate C to either A or B directly in round one. The algorithm then continues to run until no districts remain to be aggregated given the threshold \bar{C} . Prior to aggregating a given origin district to a given destination district, the algorithm checks that in cases where a district could be aggregated to several destinations, it is, in fact, uniquely added to the one to which it sends the most workers. When commuting flows between two districts are above \bar{C} in both directions, the algorithm also ensures that the smaller district is aggregated to the larger (see, Duranton, 2015a, p 184).

Based on his own application of the algorithm to Colombian municipalities, Duranton (2015a) notes that, given the gravitational nature of commuting, the algorithm's "preferred" threshold is likely to fall in the sizes of the underlying sub-national units being aggregated into metro areas.

Approach #2: The Agglomeration Index (AI)

The Agglomeration Index (AI) was originally developed by Uchida and Nelson (2009) for the World Bank's 2009 World Development Report (WDR) and, since then, has been further applied in other World Bank reports, including in World Bank and IMF (2013) and Ellis and Roberts (2016). The AI was designed by Uchida and Nelson for global application. Given the absence of O-D commuting flow data for many countries – especially developing countries – it instead relies on estimated travel times to a set of pre-defined cores to delineate the extents of metro areas. Cores are pre-defined from a global settlement point layer based on a population threshold. Rather than rely on sub-national administrative units, the AI relies on a globally gridded population data set. In such a data set, the underlying units that undergo aggregation are grid cells that are of a uniform geographic size – in practice, 30", which is approximately 1 km² at the equator.

Constructing the AI requires the specification of three thresholds – a minimum population threshold to identify settlement points that qualify as metro cores (\bar{P}), a travel time threshold (\bar{T}), and a population density threshold (\bar{D}). While Uchida and Nelson (2009) experimented with a range of thresholds, as we also do in our application to Indonesia below, the AI has become synonymous with thresholds of $\bar{P} = 50,000$, $\bar{T} = 60$ min, and $\bar{D} = 150$ people per km². Hence, for these thresholds, the AI defines a group of grid cells as constituting a metro area if each of those

cells has a population density of at least 150 people per km² and fall within a 60-min travel time radius of a settlement point that has an associated population of at least 50,000. An important feature of the AI is that, in contrast to the other two non-commuting data approaches that we discuss below, it does not include a contiguity requirement. This means that, in principle, and similar to Duranton's (2015a) commuting algorithm, a metro area does not necessarily have to consist of a single contiguous block of grid cells. Another important feature of the AI is that if there are two or more cores that fall within 60 min travel time of each other, then they merge into a single extended metro area.

Approach #3: The cluster algorithm

Rather than attempting to delineate a metro area based on its functional area, using either commuting flow data or estimated travel times, Dijkstra and Poelman's (2014) cluster algorithm simply identifies a metro area as a dense population cluster. The algorithm was originally developed with the European Union in mind but has since been applied globally. It forms the basis of the "degree of urbanization" approach that Dijkstra et al. (2019) apply globally. Given the simplicity of its data requirements, this algorithm has emerged as the preferred algorithm of not only the European Commission, but also of a coalition of international agencies that further includes the Food and Agricultural Organization (FAO), the International Labor Organization (ILO), the OECD and the World Bank. In March 2020 it was also endorsed by the United Nations' Statistical Commission as a recommended method of making international comparisons of urban areas. As with the AI, the cluster algorithm relies on a gridded population data set as input, where, in practice, this data set again has a resolution of 30". Given this data, it identifies a spatially contiguous set of grid cells as a metro if each of those grid cells satisfies a population density threshold, \bar{D}_C , and, collectively, the population of the grid cells exceeds a population threshold, \bar{P}_C .

In this paper, we will experiment with different sets of thresholds. However, in its global application, the cluster algorithm is associated with two different sets of thresholds. The first set is $\bar{D}_C = 300$ people per km² and $\bar{P}_C = 5,000$ with the resultant areas that are delineated being referred to as "Urban Clusters". Meanwhile, the second set of thresholds is $\bar{D}_C = 1,500$ people per km² and $\bar{P}_C = 50,000$ with the areas that result being labelled "Urban Centers" (Dijkstra and Poelman, 2014). Together, urban clusters and urban centers correspond to "level 1" of Dijkstra et al.'s (2019) degree of urbanization approach.

Approach #4: Thresholding of night-time lights (NTL) data

The use of night-time lights (NTL) data to identify metro areas, and, more generally, urban settlements, originated in the remote sensing literature with early applications including Imhoff et al. (1997), Sutton (2003), and Small et al. (2005). These, and also the more recent studies using this approach (see, for example, Zhang and Seto, 2011; Zhou et al., 2015; Baragwanath et al., 2020; or Dingel et al., 2020), have invariably relied on data products derived by the Colorado School of Mines' Earth Observation Group (EOG), which was formerly based at the National Oceanic and Atmospheric Association (NOAA), from sensors (Optical Line Scanner, or OLS, sensors) on-board the Defense Meteorological Satellite Program (DMSP) constellation of satellites. The derived DMSP-OLS data products cover the entire globe and are available at a resolution of 30". One deficiency of DMSP-OLS NTL data, however, is that it suffers from a well-documented "overglow" or "blooming" problem, whereby the light emitted from a given point on the earth is recorded as covering an area that extends beyond that point (see Doll, 2008, for an early discussion). This creates a tendency for the lit area of a metro to overstate its "true" extent – for example, the Pacific Ocean can be lit up as far as 50 km from the coastline near Los Angeles in the data (Pinkovskiy, 2013). The most common approach to dealing with this overglow problem has been to threshold the NTL data, considering only pixels in the data that exceed a certain luminosity, or digital number (DN), value as part of the area of a city or metro (see, for example, Imhoff et al., 1997; Small et al., 2005; Zhou et al., 2015; Ellis and Roberts, 2016; Baragwanath et al., 2020; Dingel et al., 2020). A con-

tiguous cluster of grid cells that falls above the applied threshold is then classified as being a "metro" area.^{4,5}

More recently, however, DMSP-OLS NTL data has been superseded by NTL data collected from a newer satellite sensor, the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor, launched in 2011. This sensor collects NTL data at a far higher resolution than the old DMSP-OLS sensors and the derived data products are also much less affected by overglow (Small, 2019). We use the new VIIRS data to delineate metro areas. Specifically, we use EOG's recently released 2015 annual composite product.⁶ This product reports luminosity values, calculated as an annual average over all cloud-free nights in 2015, at a resolution of 15", which is equal to 460 m² at the equator. Prior to averaging, EOG applies filtering techniques to remove data that is affected by stray light, lightning, and lunar illumination. EOG likewise filters out lights from aurora, fires, boats and other temporary lights. Although the VIIRS NTL data is much less affected by overglow, it, nevertheless, records light emitted to the nighttime sky by all human activities, including light at very low levels outside of what may be considered metro or urban areas (for example, light emitted by traffic along roads connecting different cities and villages). For this reason, the use of a threshold may still be required to properly delineate metro areas. As with papers that have used the DMSP-OLS data for the same purpose, we consider a contiguous cluster of grid cells that falls above any imposed threshold as representing a potential "metro" area, where we experiment with a range of different thresholds in this paper.

3. Data

3.1. Data sources

The data that we use to apply the four approaches to delineating metro areas in Indonesia come from several sources. For the commuting algorithm, we use data on O-D commuting flows between all 497 Indonesian districts, as defined by their 2013 boundaries, that we derive from the August rounds of Indonesia's national labor force survey (*Survei Angkatan Kerja Nasional, SAKERNAS*) for the years 2013–2015.⁷ In doing so, we measure the commuting flow from a given origin district i to a destination district j as the share of workers who live in i but commute daily to work in j , where – following SAKERNAS – workers are defined to include all employed wage workers including casual workers, self-employed (or "own account") workers, and unpaid family workers, where anyone who worked for at least one hour consecutively in the previous week is considered employed. We also use SAKERNAS data, from the August 2014 round, as our main source of data on nominal wages and a host of individual worker characteristics, in the estimation of the agglomeration wage premium in Sections 5 and 6.

Both the AI and the cluster algorithm require a gridded population data set. We use the *Landsat-2012* gridded population data set produced by Oak Ridge National Laboratory. This population grid has a resolution of 30". The grid is derived through distributing population data for sub-national administrative units across grid cells using a

⁴ In addition to a luminosity threshold, some studies apply an overall population threshold analogous to that used by the cluster algorithm. For example, Dingel et al. (2020) use a population threshold of 100,000.

⁵ Rather than using a luminosity threshold to address the overglow problem in delineating metro areas using DMSP-OLS NTL data, Ch et al. (2020) instead apply an algorithm for correcting overglow developed by Abrahams et al. (2018). This algorithm is not suitable for application to the VIIRS nighttime lights data that we utilize in this paper.

⁶ This product is available for download from: https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html or https://eodata.mines.edu/download_dnb_composites.html.

⁷ Sampling in the SAKERNAS August rounds is stratified at the district level for these years. We average the commuting flows over three years to smooth-out any temporary measurement error in the survey data.

modelling process that relies on other geo-spatial data sources and high-resolution satellite imagery analysis.⁸ The *Landscan* grid is one of the most established global population grids and has been widely used in social scientific research. This includes by Henderson et al. (2020), who use the data in identifying urban areas and for constructing measures of population and economic density for six African countries. Importantly, Henderson et al. (2020) “ground-truth” the *Landscan* data, reaching the conclusion that it does a good job in estimating population at a fine spatial scale.

In addition to gridded population data, the AI also requires data on estimated travel times and settlement points. The travel time data originally used by Uchida and Nelson (2009) for the AI was based on “... estimates of the time required to travel 1 km over different road and off-road surfaces...” and was derived from a cost surface that was constructed from a variety of Geographic Information System (GIS) data layers. These layers included data on road and rail networks, navigable rivers and water bodies, travel delays for crossing international borders, roughness of terrain, and foot-based travel for off-road and paths. We use an updated version of this same cost surface layer from Berg et al. (2017), which is derived from more recent (i.e. circa 2010 as opposed to circa 2000) data on roads, railroads, and land cover. Meanwhile, for data on settlement points, we use version 1.01 of CIESIN's (2017) Global Rural – Urban Mapping Project (GRUMP) Settlement Point layer.

Finally, as already described in Section 2, the NTL data that we use is VIIRS NTL data taken from EOG's 2015 annual composite product.

3.2. Mapping to districts

One issue that we face in generating results that can be compared across the different metro area delineation approaches is that while the commuting algorithm uses sub-national administrative units – in our case, Indonesian districts – as the “building blocks” for metro areas, the remaining approaches rely on much higher resolution gridded data sets. This means that while the perimeters of the metro areas defined by the commuting algorithm are constrained to follow district boundaries, this is not the case for the metro areas generated by the other approaches.

This is, in principle, a highly attractive feature of the AI, cluster and NTL approaches. But, importantly, it poses difficulties for any empirical analysis that wishes to use the metro areas based on these approaches. This is because other data that the researcher might wish to match to these metro areas will often only be available for sub-national administrative areas or, in the case of household and firm survey micro-data, include location identifiers for such areas only. Furthermore, household or firm survey micro-data are typically obtained using a random sampling strategy stratified at the level of sub-national administrative areas, so that it is not representative at a finer level of spatial aggregation even if more spatially disaggregated location identifiers are provided.

Given the above, we need to map the urban extents generated by the AI, cluster, and NTL approaches to (aggregations of) Indonesian districts.⁹ We do this by always applying the same basic rule: we associate two or more districts with a single urban extent if at least 50% of each

⁸ See http://web.ornl.gov/sci/landscan/landscan_documentation.shtml for more information.

⁹ The median area of an Indonesian district is 1943 km² with a minimum of 9.6 km² and a maximum of 44,177 km². However, 76.4% of Indonesia's urban population in 2014 lived in districts of below median area, while 53% lived in districts of area less than 1000 km². Despite this still substantial size, the metro areas generated using the different approaches map very well to these administrative districts on Java-Bali where urbanization in Indonesia is most advanced (see Fig. 2 in Section 4.1 below). For the rest of Indonesia, where districts are more likely to have rural parts, the mapping is not as tight. This will be important when it comes to explaining the difference in the behavior of the estimated agglomeration wage premium to the strictness of the definition of metro areas in Section 5.3 of the paper.

district's population belongs to that extent. In this way, we aggregate districts to construct approximations of the urban extents implied by each non-commuting data approach.¹⁰ Analogous to the commuting algorithm, we only consider a given urban extent generated by each of the AI, cluster and NTL approaches to represent a metro area if that extent maps to two or more districts. Any urban extent that fully lies within an administrative district's boundaries is not considered to be a metro area. As such, all four approaches make a, possibly different, distinction between, on the one hand, workers who are employed in urban areas that span multiple districts (metropolitan areas) and, on the other hand, workers who are employed in urban areas that are simply confined within the boundaries of a single administrative district.

In Section 4, we mostly focus on the difference between the four approaches in terms of the number, size, and spatial extent of the (multi-district) metropolitan areas that they generate. But, importantly, in Section 5, and analogous to Chauvin et al. (2017) in the case of India, we always include all urban workers, including those who work outside the metropolitan areas, in our regressions that estimate the agglomeration wage premium. In these regressions, each urban worker is either associated with (the density of) the single administrative district or the (multi-district) metropolitan area in which (s)he works.

4. The (ir)relevance of metro definitions: Indonesia's metro area landscape

We first document the important differences between the four delineation approaches in terms of the metro area landscapes that they generate at different threshold(s). We structure our discussion by always explicitly distinguishing between what we find on Java-Bali and in the rest of Indonesia. As mentioned before, Java-Bali is very different from the rest of Indonesia in its (very) high population density, which, as we will see, has important consequences for each approach's ability to generate plausible depictions of the metro area landscape.

4.1. Key descriptive characteristics of identified metropolitan areas

Fig. 1 summarizes how three key statistics – the total number of metro areas identified, the mean number of districts per identified metro area, and the share of the urban population living in the identified metros – depend on the selected threshold(s) for each of the four approaches. Table A1.1 in Appendix 1 presents further statistics for the metro areas delineated by each of the four approaches at different thresholds, including the identity of the largest metro area identified. For each approach, the graphs in Fig. 1 show the effects of moving from a more relaxed threshold (or, in the cases of the AI and cluster approaches, set of thresholds) to a stricter threshold (or set of thresholds), looking across the x-axis from left to right.¹¹ They do so for Indonesia as a whole as well as for Java-Bali and the rest of Indonesia separately. Fig. 2 then complements Fig. 1 by mapping the metro areas on Java-Bali for selected thresholds for each of the four approaches. For the AI, this includes the conventional thresholds that have been taken as globally most applicable. It also includes the “urban cluster” and “urban center” thresholds for the cluster algorithm (see Section 2 above).

From Figs. 1 and 2, as well as Table A1.1 in Appendix 1, clear differences emerge between the metro areas identified by the commuting

¹⁰ We have experimented with the sensitivity of our results to the 50% threshold by increasing it, in steps of 5 percentage points, up to 80%. Using a higher threshold does not qualitatively affect any of our main findings.

¹¹ For both the AI and the cluster algorithms, the population density threshold is more important in driving the descriptive results than the (core/overall) population threshold. This is particularly the case for the cluster algorithm, where, for a given population density threshold, both the overall number and sizes of the metros identified remain almost invariant to the population threshold used.

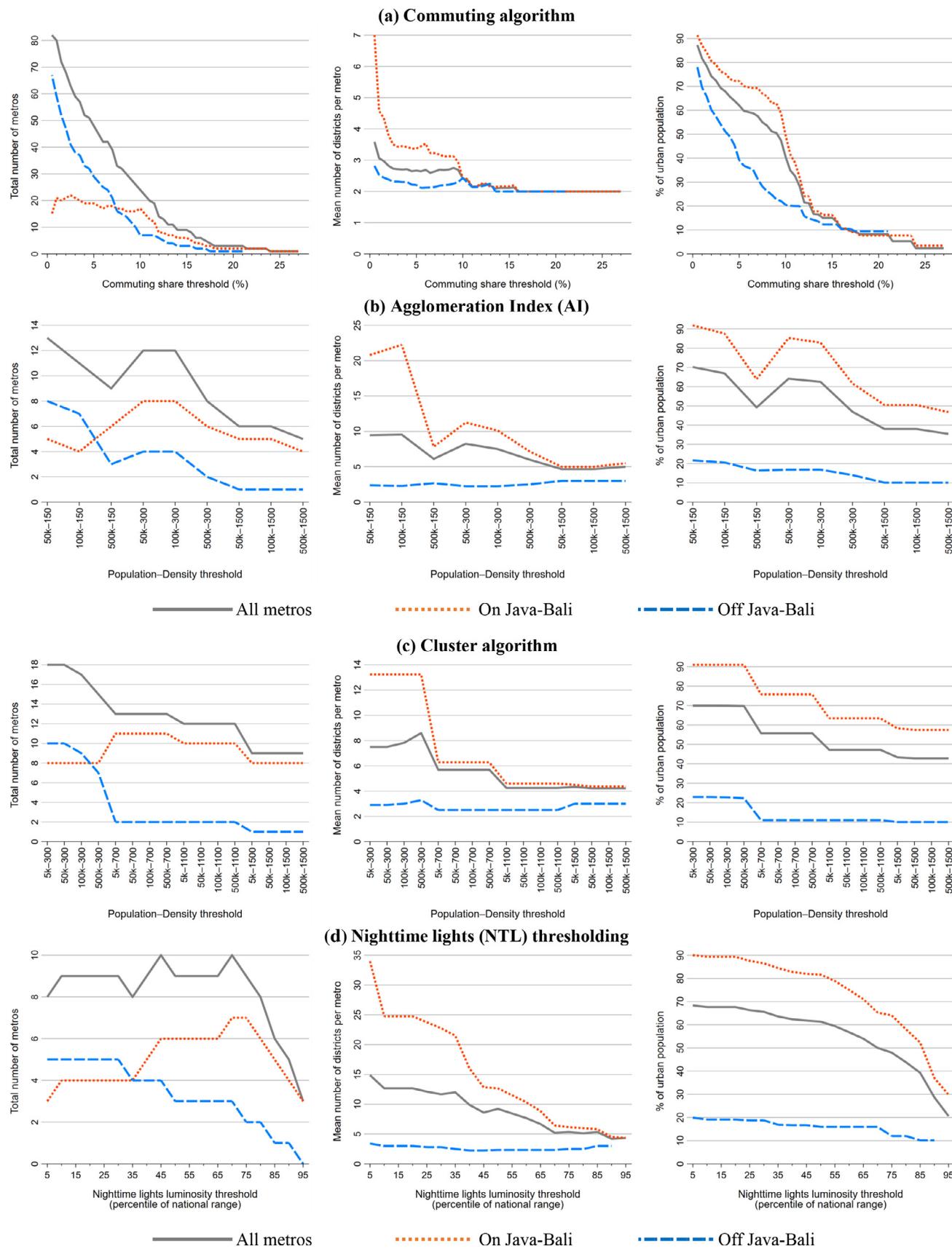


Fig. 1. Total number of metros, mean number of districts per metro, and the percentage of urban population in metros.

Notes: $k = 1000$. For on and off Java-Bali, the percentages of urban population in metros were calculated out of the total urban population on the corresponding island-regions. For the Agglomeration Index (AI), the travel time threshold is held fixed at 60 min.

Source: See Section 3.1. Urban population data from Indonesia's 2014 national socio-economic survey (*Survei Sosial Ekonomi Nasional*, SUSENAS).

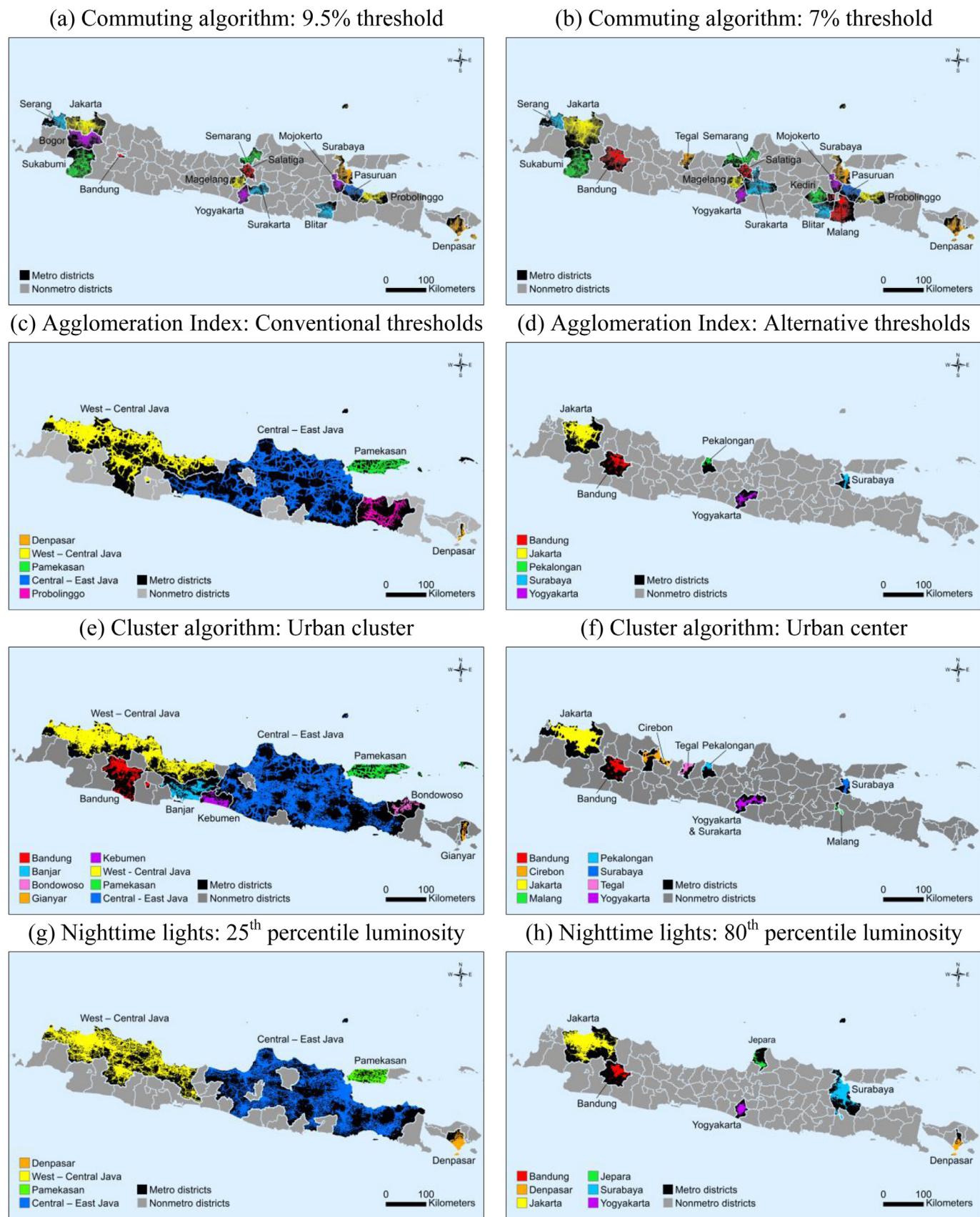


Fig. 2. Metro areas defined by several selected approaches/thresholds (Java-Bali only).

Notes: In panel (c), the thresholds are $\bar{P} = 50,000$, $\bar{D} = 150$ people per km^2 , and $\bar{T} = 60$ min, while in panel (d) they are $\bar{P} = 50,000$, $\bar{D} = 1,500$ people per km^2 , and $\bar{T} = 60$ min. The panel (e) and (f) thresholds are $\bar{D}_C = 300$ people per km^2 and $\bar{P}_C = 5,000$, and $\bar{D}_C = 1,500$ people per km^2 and $\bar{P}_C = 50,000$, respectively.

Source: See Section 3.1

algorithm on the one hand and the non-commuting data approaches on the other. This is the case both on and off Java-Bali.

On Java-Bali. For Java-Bali, the commuting algorithm steadily adds, relatively small, metro areas as its threshold is relaxed (looking across the x-axis from right to left), starting from a situation where only a (very) small proportion of Java-Bali's urban population is classified as living in metro areas. At commuting flow thresholds below 9% it identifies up to 21 metro areas, each consisting of only a few districts, covering more than 60% of Java-Bali's urban population. By contrast, the other approaches always capture a relatively large share of Java-Bali's urban population even at the strictest thresholds considered. They also show a similar tendency for metro areas to grow in both number and size initially, as these thresholds are relaxed. But, as the sizes of the metro areas grow these metros themselves start to merge together. As the thresholds pass a certain level of looseness, the number of metros begins to fall, with each metro, on average, consisting of an increasing number of districts. As a result, almost the whole of Java-Bali becomes identified as consisting of a few, extremely large, metro areas (see also parts (c), (e) and (g) of Fig. 2 for the AI, cluster and NTL approaches respectively). Interestingly, Fig. 2(c) corresponds to the global thresholds that have become synonymous with the AI, while Fig. 2(e) corresponds to the "urban cluster" set of thresholds associated with the cluster algorithm. Clearly, these thresholds are inappropriate for Java-Bali, and, presumably, therefore, also for other densely populated parts of the world such as Bangladesh, Northeast India or the Nile delta.

Off Java-Bali. Off Java-Bali, the four approaches show rather different properties to those which they display for Java-Bali with sharp differences, again, between the commuting algorithm on the one hand and the remaining approaches on the other. The commuting algorithm only identifies a relatively small number of metro areas at commuting flow thresholds of 10% and above. However, as its threshold falls below 10%, the number of metro areas rapidly increases, exceeding the number identified on Java-Bali at a threshold of 7%, and reaching more than 30 (40) for thresholds below 5 (2.5%). These metros are comprised, on average, of only a little over two districts, compared to an average of 3–4 districts per metro area on Java-Bali at thresholds below 10%. Also, the percentage of the urban population living in the identified metro areas is always smaller compared to that found on Java-Bali. This percentage is, nevertheless, much larger than for any of the other (non-commuting data) approaches. It is only when using extremely low population density, in the cases of the AI and cluster algorithms, or NTL/luminosity thresholds that we observe any significant (growth in the) number of metro areas at all. Even at the most relaxed thresholds, the number of metros identified off Java-Bali is far fewer than the number identified by the commuting algorithm at commuting flow thresholds below 10%. Moreover, and in contrast to what we saw for Java-Bali when using either of the three non-commuting based approaches, these few metros are each formed by a small number of underlying administrative districts.

Fig. 3 provides further insights into the differences in results, both on and off Java-Bali, between the commuting algorithm and the three other approaches. For selected metros at selected thresholds, this figure shows the level of agreement, in terms of the districts identified as forming the metro, between the different approaches. In the case of Java-Bali's two largest cities, Jakarta and Surabaya, all approaches agree on a specific sub-set of core districts. However, levels of agreement decline as we move further away from these core districts. This happens because approaches with stricter thresholds (i.e. higher commuting flow, population density or luminosity thresholds) tend to classify as metro districts only sub-sets of those districts that are classified as being part of a metro area under approaches with more relaxed thresholds. In general, it is the AI with its conventional thresholds which results in the largest metro areas.

For smaller metros, typically found off Java-Bali, such as Makassar and Manado on Sulawesi shown in Fig. 3, the picture is very different. While the commuting algorithm identifies these as metros based on the

strength of cross-district commuting flows, they are only identified by the non-commuting data approaches at lower density or NTL/luminosity thresholds, if at all. These examples illustrate a more general point: compared to the commuting algorithm, all three non-commuting data approaches require high density or NTL/luminosity thresholds to avoid producing metro areas that are "too large" on Java-Bali, where average population density is high. However, these same high thresholds lead these approaches to, in many cases, completely miss metro areas identified by the commuting algorithm off Java-Bali where average population density is much lower. This raises questions about the possibility to define metro areas around the world using a single, globally applicable, approach based on a single, globally applicable, (set of) threshold(s).

4.2. Jaccard indices – agreement on the identified metro districts across approaches/thresholds

Additional information on the overall spatial level of agreement between the different approaches is provided by Fig. 4, which presents values of the Jaccard index for pairwise comparisons of the "metro maps" generated by the different approaches for selected "reasonable" thresholds – i.e. thresholds that do not lead to the whole of Java-Bali being covered by a few unrealistically large metro areas only.¹² In Fig. 4(a), the Jaccard index measures the proportion of aggregate land area of metro districts in two maps among the aggregate land area of metro districts in either of the maps. More formally, if we denote the set of districts that are classified as metro districts in one map by A and the set of districts that are classified as metro districts in a second map by B then the Jaccard index is given by $J(A, B) = \text{Area}_{|A \cap B|} / \text{Area}_{|A \cup B|}$. Meanwhile, in Fig. 4(b), the Jaccard index is instead calculated as the aggregate urban population of districts that belong to $A \cap B$ divided by the aggregate urban population of districts that belong to $A \cup B$.¹³

Despite the difficulties associated with the Jaccard index,¹⁴ Fig. 4 provides several interesting insights. First, the level of agreement between any pair of maps tends to be higher when comparing urban population than when comparing land area. This is consistent with more peripheral districts, on which the different approaches tend to disagree more, having lower average population densities. Second, when comparing maps generated by different approaches, levels of agreement between the commuting algorithm and the non-commuting data approaches are lower than those between the non-commuting data approaches themselves, irrespective of whether we are comparing land area or urban population. Third, these differences in levels of agreement are particularly marked off Java-Bali. This is because, on Java-Bali, relatively large numbers of districts are classified as belonging to metro areas for all approaches, even though the commuting algorithm tends to split these districts up into many more metro areas than the other approaches. By contrast, off Java-Bali, where population densities are much lower, the commuting algorithm identifies 10 and 21 metro areas, comprising of 23 and 45 districts in total, for commuting flow thresholds of 9.5 and

¹² Jaccard index results based on a more complete set of thresholds are available upon request.

¹³ Neither the area- nor the population-based Jaccard index as calculated here capture whether a pair of maps depict a given district as belonging to the same metro area in both maps. Rather, the indices just capture whether a given district belongs to some metro area in both maps – i.e. in the terminology of de Bellefon et al. (2020), the Jaccard indices presented here are "urban Jaccard" indices rather than "city Jaccard" indices.

¹⁴ These difficulties are that: (1) no clear guidelines exist as to what values of the index correspond to a "strong" versus "weak" level of agreement, and (2) it is not possible to formally test whether differences in the value of the index between, for example, two different pairwise comparisons are statistically significant. de Bellefon et al. (2020) have recently introduced methods to address the second difficulty.

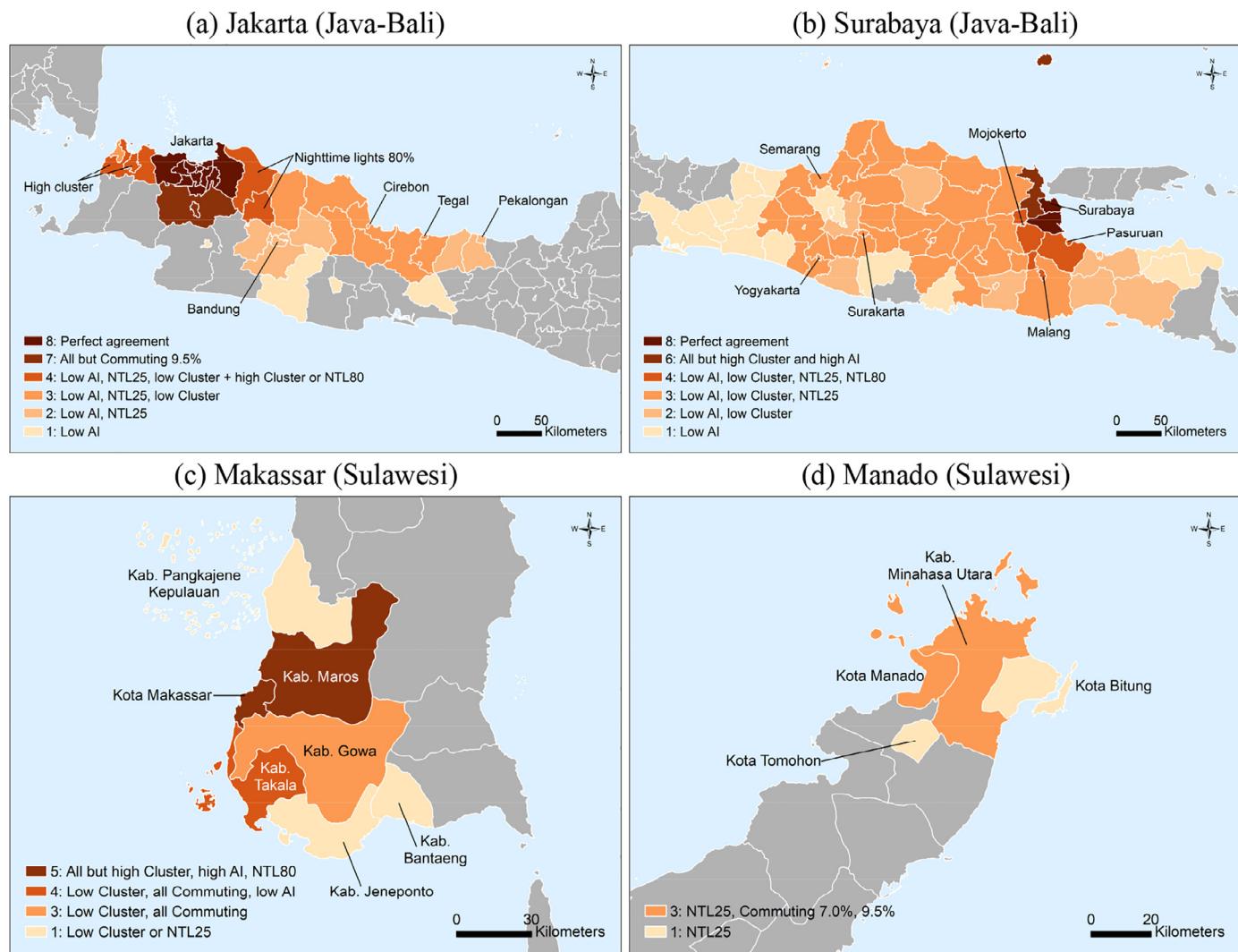


Fig. 3. Visualization of level of agreement in component districts for specific metro areas

Notes: Each map compares the four metro delineation approaches for two different sets of thresholds – a strict set of thresholds and a loose set of thresholds. Hence, each map compares a total of eight different delineations of metro areas (4 approaches \times 2 sets of thresholds). Districts are color-coded and numbered according to the number of these eight delineations that identify them as belonging to a specific metro area. Commuting 9.5% and Commuting 7.0% refer to the commuting algorithm with commuting flow thresholds of 9.5% and 7% respectively; low AI and high AI refer to the AI with thresholds of $D = 150$ people per km^2 and $P = 50,000$, and $D = 1,500$ people per km^2 and $P = 50,000$, respectively (and a travel time of 60 min in both cases); low Cluster and high Cluster refer to the cluster algorithm with the “urban cluster” and “urban center” sets of thresholds respectively; and NTL25 and NTL80 refer to the NTL approach with luminosity thresholds at the 25th and 80th percentiles of Indonesia’s national distribution of luminosity values respectively. Kab. denotes Kabupaten.

Source: See Section 3.1

7% respectively, whereas the non-commuting data approaches always identify extremely few districts as belonging to metro areas.

In sum, we find marked differences in the metropolitan area landscape generated by the different approaches and thresholds used to delineate metro areas. Two things stand out. First, we find the largest, most salient, differences between the metro area landscapes generated by the commuting algorithm and the non-commuting data approaches. Second, this is the case throughout Indonesia, but for very different reasons on densely populated Java-Bali compared to the much less densely populated rest of Indonesia. On Java-Bali, the use of the non-commuting data approaches very easily results in identifying only a few, very large metro areas, except at the highest population density and luminosity thresholds used. Using a commuting approach instead results in many more metro areas, each covering far fewer districts. By contrast, off Java-Bali, the use of the non-commuting data approaches results in missing out on many metro areas, i.e. not classifying districts as part of the same metro area that would, in fact, be classified as such based on the strength of the commuting flows between them.

5. The (ir)relevance of metro definitions to estimation of the agglomeration wage premium

Given the marked differences in the metropolitan area landscape generated by the different approaches and thresholds, a natural next question is whether the choice of metro definition also matters when estimating key empirical relationships in urban economics. In this section, we investigate this question for one such key empirical relationship – namely, the existence of a significant positive relationship between the nominal wage that an individual worker earns and the density of the urban area in which (s)he works. The presence (size) of such an **agglomeration, or density, wage premium** is taken as evidence of (the strength of) the productivity benefits associated with density.¹⁵ In particular, we show

¹⁵ Our focus on a measure of density follows Ciccone and Hall (1996), De La Roca and Puga (2017), Quintero and Roberts (2018) and Henderson et al. (2020), among others. As Duranton and Puga (2020) note,

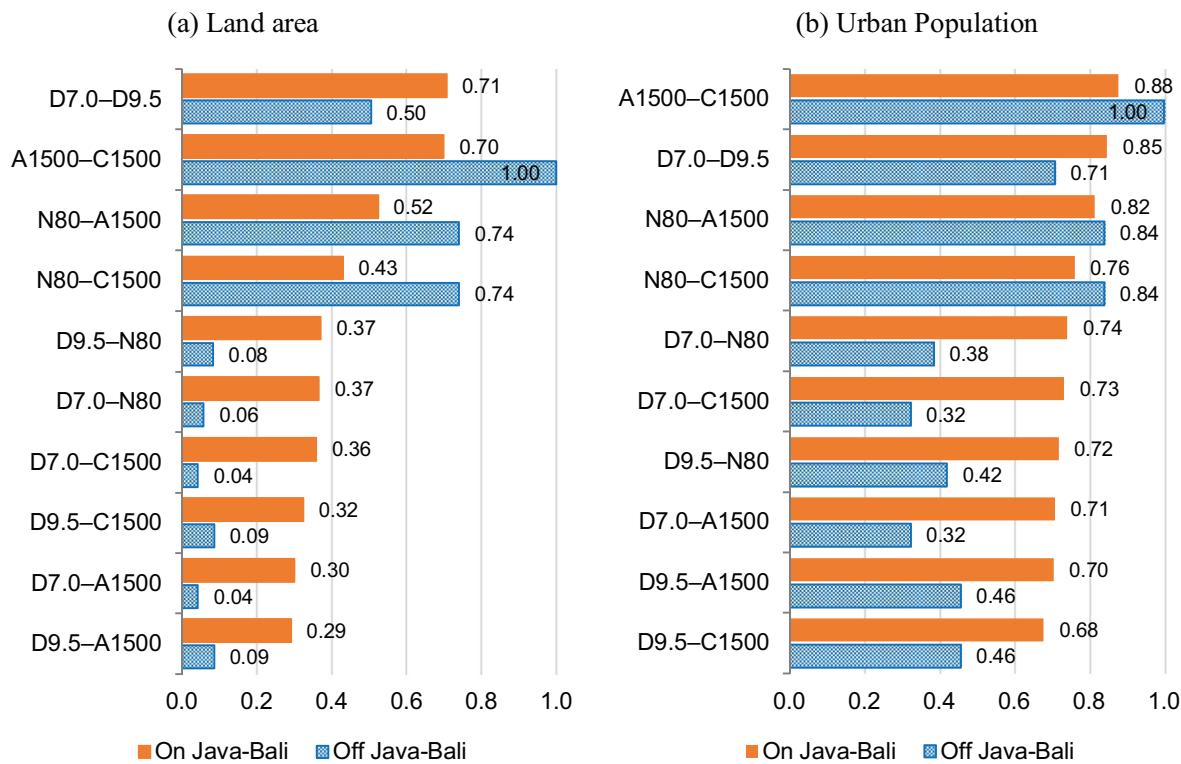


Fig. 4. Jaccard indices for pairwise map comparisons – selected thresholds

Notes: D7.0 and D9.5 refer to Duranton's commuting algorithm with a commuting flow threshold of 7.0% and 9.5% respectively; A1500 to the Agglomeration Index with thresholds $\bar{P} = 50,000$, $\bar{D} = 1,500$ people per km^2 , and $\bar{T} = 60$ min; C1500 to the cluster algorithm with thresholds $\bar{D}_C = 1,500$ people per km^2 and $\bar{P}_C = 50,000$; and N80 to the NTL approach with a threshold at the 80th percentile of the distribution of luminosity values for Indonesia. In both panels, charts are sorted in descending order of index values for Java-Bali.

how the existence and estimated size of this premium for both Java-Bali and the rest of Indonesia depends on the choice of approach to delineating metro areas and the accompanying choice of threshold(s). In doing so, we retain all urban workers in the sample, so that all our regressions have the same sample size. As such, the choice of approach and threshold(s) affects whether a given worker is classified as being employed in either a metro or in a non-metro urban area. When a worker is classified as being employed in a metro area, it also potentially affects which metro area the worker is associated with, as well as the measured density of that area. Following the presentation of the results for the estimated agglomeration wage premium, we provide a simple explanation for the differences in results between different thresholds for a given approach to defining metro areas.

5.1. Estimating the agglomeration wage premium in Indonesia

While there exists a large and well-established literature that empirically examines the size of the agglomeration wage premium for developed countries (for excellent reviews, see Rosenthal and Strange, 2004; Combes and Gobillon, 2015; Ahfeldt and Pietrostefani, 2019), there are relatively few papers that do so for developing countries. The main exceptions include recent papers by

the correlation between a city's density and its population tends to be very tight provided density is measured with sufficient care, while the theoretical literature is largely agnostic as to whether a measure of density or scale is the "correct" measure of agglomeration. Regression results when instead using total population as the agglomeration variable are available upon request. Briant et al. (2010) and Holmes and Lee (2010) also show how the measured degree of spatial concentration and the evidence on Zipf's Law can depend on the choice of spatial units used.

Duranton (2016), Chauvin et al. (2017), Quintero and Roberts (2018), Roberts (2018b), and Henderson et al. (2020) discussed in Section 1.¹⁶ These papers use detailed geospatial micro datasets at the individual worker level drawn from either household or labor force surveys, and estimate the size of the agglomeration wage premium by regressing an individual worker's nominal wage on a measure of the density of the urban or metropolitan area in which (s)he is employed while controlling for a host of observable worker characteristics.

We follow a similar approach in estimating Indonesia's agglomeration wage premium. We draw on rich micro-data for Indonesian workers from the same labor force survey – the SAKERNAS – that we used to measure commuting flows (see Section 3.1), and estimate the agglomeration wage premium using the following basic regression:

$$w_{ijd} = \alpha_j + X_1 \gamma_1 + X_d \gamma_2 + \beta S_d^m + \epsilon_{ijd} \quad (1)$$

where our dependent variable w_{ijd} denotes the (natural log) hourly nominal wage of individual i , working in the urban part of district d , in industry j .¹⁷ α_j denotes a full set of 186 3-digit industry fixed effects (with industries as defined in the 2000 Indonesian Standard Industrial Classi-

¹⁶ See also Combes et al. (2015) and Combes et al. (2019). They examine agglomeration effects for China but use a different empirical framework to the papers cited in the main text.

¹⁷ SAKERNAS reports monthly income for sampled individuals. Based on this, and a person's reported total working hours during the last week, we calculate his/her hourly wage as: (monthly income / (365/12)) × (7/hours worked last week). SAKERNAS reports both monthly incomes earned in cash and in goods. In all results reported here, we use total monthly income in cash and in goods combined. All our findings are robust to using only total monthly income in cash (results available upon request).

fication, KBLI 2000).¹⁸ \mathbf{X}_i denotes a vector of individual worker characteristics. Specifically, we control for a worker's age and its square, which can be considered to proxy for his/her *overall workforce experience*. We also include a worker's *experience on his/her current job* and its square, which is measured as the number of years (s)he has been working on the job, as well as a dummy variable indicating whether (s)he ever held a job before his/her current job. Next, we control for a worker's highest completed level of education,¹⁹ as well as two dummy variables indicating whether (s)he completed at least one, respectively two, additional formal training courses. Finally, we include a dummy variable indicating whether a worker is an "own account worker" (i.e. self-employed) or an employee. \mathbf{X}_d is a set of eight dummy variables for the main Indonesian islands (island-groups) on which districts are located (Bali, Java, Kalimantan, Maluku, Nusa Tenggara, Papua, Sulawesi, and Sumatra), and ε_{yd} captures all other unobserved nominal wage determinants.

S_d^m is our main independent density variable of interest. We follow De La Roca and Puga (2017) in the specification of this variable. For all urban workers employed in districts that belong to a (multi-district) metro area, S_d^m is thus defined as "*the (ln) number of people within 10 km of the average person in the metropolitan area, m, to which district, d, belongs*", i.e. it takes the same value for all workers employed in the same metro area. For all urban workers employed in districts that do not belong to a metro area, it is simply defined as "*the (ln) number of people within 10 km of the average person in the district, d*", which is analogous to treating a district as being its own metro area. Duranton and Puga (2020) refer to this particular measure of density as "experienced density" to contrast it with the more naïve measure of density that is derived by simply dividing an area's population by its land area. As they highlight, the experienced density measure helps deal with the uneven tightness of area boundaries and, compared to "naïve density", better captures how close the typical person is to other people when the population is unevenly distributed. We calculate experienced density using the same Landscan-2012 gridded population data that we used to define metro areas using the AI and cluster algorithms (see Section 3.1 above). One advantage of Landscan in this context is that, unlike other available gridded population data sets such as GHS-POP (Schiavina et al., 2019) and WorldPop (Tatem, 2017), it is an ambient, as opposed to residential, population grid. Hence, it is intended to capture where people are located, on average, during the day, which will be affected by where they work, as opposed to just where they reside. For the purposes of estimating the agglomeration wage premium, this leads to a more relevant measure of density than residential population density.

Finally, β is our main coefficient of interest. It measures Indonesia's agglomeration wage premium – i.e. the elasticity of the nominal wage with respect to density. This coefficient is of interest in its own right, but in this section, we are particularly interested in how the estimated size (and significance) of β depends on the approach and thresholds used to define metro areas.

In estimating equation [1], we follow De La Roca and Puga (2017) as closely as possible, restricting attention to working age (15–64 years old) males that reported a non-zero income in the August 2014 round of SAKERNAS. This initial sample consists of 116,156 workers. Next, we further exclude workers employed in agriculture, forestry, livestock and fishing, mining and quarrying, public administration and defense, education, health and social work. These activities are typically rural or much more controlled / regulated by Indonesia's national and local governments. Also, we only focus on workers who have worked at least 32 hours (four working days) during the last week. These restrictions

¹⁸ Results, available upon request, are very similar when including either no or 88 two-digit sector dummies instead.

¹⁹ SAKERNAS distinguishes 14 different levels of educational attainment. From lowest to highest, these levels are: no schooling, incomplete primary school, primary school, package A, general junior high school, vocational junior high school, package B, general senior high school, vocational senior high school, package C, diploma I/II, diploma III, div/S1, and S2/S3 (graduate level).

reduce the sample to 56,577 workers. Finally, we restrict the sample to urban workers only, leading to a final sample of 37,869 workers. The SAKERNAS classifies each worker as living in an urban or rural area.²⁰ And, since we only know the location of workers at the district level, this focus on urban workers ensures that we only include those workers that are most likely to live and work in the urban parts of a district.²¹

5.2. Main findings

Fig. 5 shows our main results. Based on our findings in Section 4, we always estimate the agglomeration wage premium for Java-Bali (panel a) and the rest of Indonesia (panel b) separately. Each panel reports the estimated agglomeration wage premium using each of the four different metro delineation approaches. For each approach, the graphs show how the estimated β changes when varying the "key" threshold.^{22,23} For both the AI and the cluster algorithms, this is the population density threshold, defined at the pixel level. In our Indonesian setting, this threshold has the most profound consequences for the number and spatial extents of the metropolitan areas that these approaches identify (see Fig. 1, and footnote 11). Dijkstra et al. (2019) likewise show that, globally, the results of the cluster approach, especially for the definition of urban centers, are more sensitive to its population density threshold than to its overall population threshold. Meanwhile, for the commuting and NTL approaches the key (only) thresholds are the commuting flow and luminosity thresholds respectively. These thresholds themselves act, in practice, as *de facto* density thresholds. Fig. A2.1 in Appendix 2 shows that, as with the use of a higher density threshold in the AI and cluster approaches, the higher the commuting flow or luminosity threshold, the fewer and denser the identified (multi-district) metro areas become. Furthermore, Fig. A2.2 in Appendix 2 shows the very strong correlation between a district's population density and its luminosity.

Finally, for comparison, the dotted lines in Fig. 5 show the estimated agglomeration wage premium when simply taking each Indonesian administrative district to be its own metro area. This is the strictest possible metro definition given the geospatial information on individual workers' job locations that is available in the SAKERNAS. All the other definitions that we use to delineate metro areas result in at least one, but often many more, (multi-district) metro areas (see Section 4).

Three things are worth highlighting based on our findings. First, the agglomeration wage premium is several orders of magnitude larger on Java-Bali than in the rest of Indonesia, regardless of the metro area de-

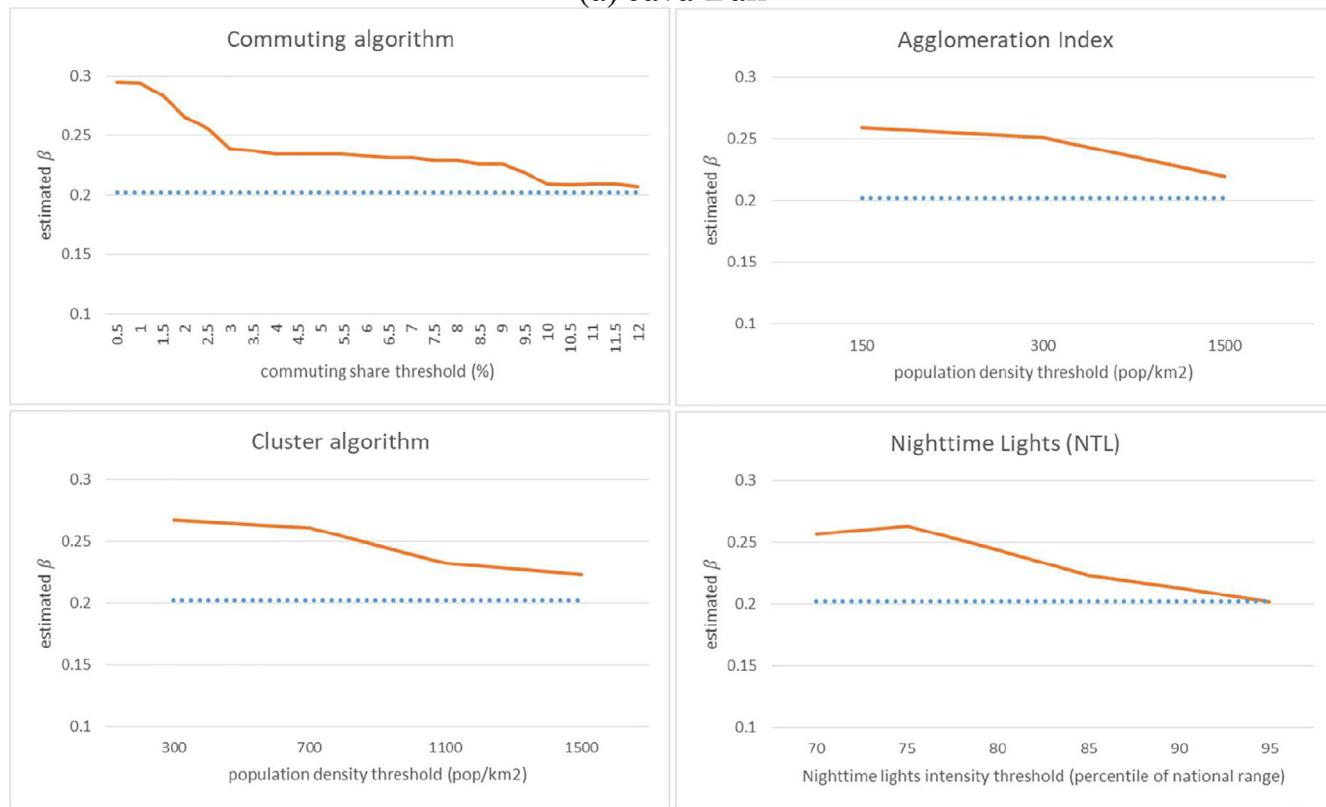
²⁰ This classification is based on a composite scoring system which assesses Indonesian villages as either urban or rural based on their possession of certain "urban characteristics" (Badan Pusat Statistik Regulation 37/2010): (i) population density; (ii) the structure of the local economy (specifically, the share of agricultural households); (iii) the percentage of households with certain types of infrastructure (i.e. electricity and telephone networks); and (iv) the availability of urban facilities (those considered are schools, hospitals, a market, shops, and a cinema, as well as recreational facilities such as a hotel, a salon, a billiard hall, a discothèque and a massage parlor).

²¹ Our main findings are robust to the sample choices outlined in this paragraph. Results when, for example, extending the sample to also include all workers living in the rural parts of districts or all female workers are available upon request.

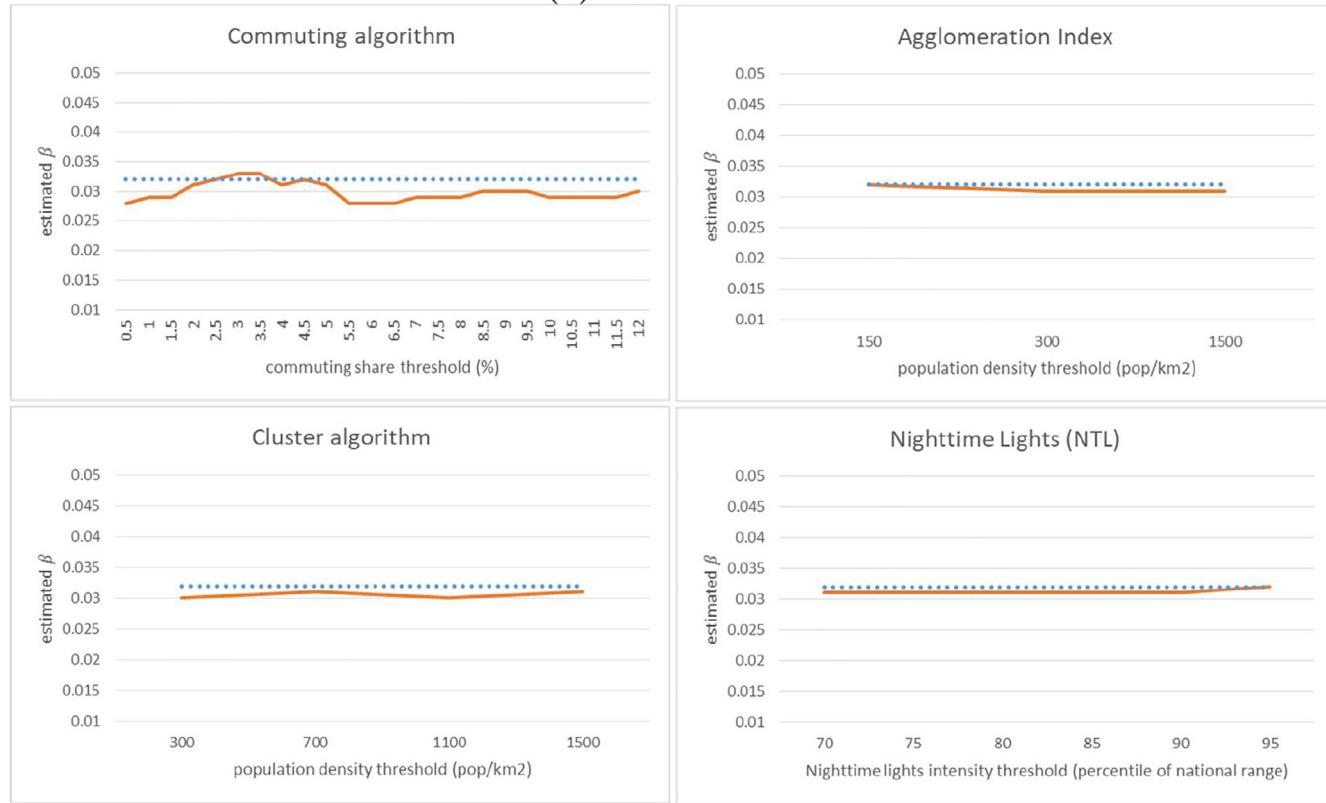
²² In the cases of the commuting and NTL approaches, Fig. 5 only shows results for truncated ranges of each approach's key threshold that generate the most plausible depictions of Indonesia's metro area landscape. A plausible depiction of Indonesia's metro area landscape is taken to be one in which: (i) Jakarta is clearly identifiable as the largest metro area; (ii) the share of the population that lives in metro areas does not exceed the share of Indonesia's population (51.3%) that was classified as urban in 2012; and/or (iii) the number of metro areas identified on Java-Bali is at least as great as the number identified off Java-Bali.

²³ We always cluster our standard errors at the metro or district level (depending on whether a district is classified as belonging to a multi-district metro area or not), i.e. at the same level at which experienced density varies.

(a) Java-Bali



(b) Off Java-Bali

**Fig. 5.** Indonesia's agglomeration wage premium, by approach and threshold(s).

Notes: All the depicted estimates of the agglomeration wage premium are significant at the 5% level. Adding confidence bands to the figures is possible but would make the patterns in the estimates less visible due to the resulting extended y-axis needed to accommodate the bands. For both the AI and cluster algorithms, the (core/overall) population threshold is kept fixed at 50,000; the travel time threshold for the AI is also kept fixed at 60 min.

lineation method used (compare panel 5(a) to panel 5(b)). On Java-Bali, a 1% increase in density is associated with a 0.2–0.3% increase in nominal wages, which is comparable to estimates found for China (Chauvin et al., 2017). Off Java-Bali this elasticity is still significant, but much smaller, only about 0.03, which is comparable to what other studies have found for countries in Latin America and the Caribbean (see Quintero and Roberts, 2018; Roberts, 2018b). We explore possible explanations for this large difference between Java-Bali and the rest of Indonesia in Section 6.

Second, the estimated agglomeration wage premium off Java-Bali hardly differs when using different approaches or threshold(s) to delineate metro areas – see Fig. 5(b). Moreover, it is always very similar to that found when simply using Indonesia's district boundaries to define metro areas. For the AI, cluster and NTL approaches, the estimated agglomeration wage premium is basically the same regardless of the density or luminosity threshold used (note the scale of the y-axes in panel b). Only when applying the commuting algorithm do we find some very modest differences in our estimates depending on the commuting flow threshold used, with estimates ranging from 0.028 to 0.033.

Third, and in sharp contrast, the estimated agglomeration wage premium for Java-Bali does depend on the approach/threshold used to define metro areas, with an elasticity that ranges between 0.202 and 0.295. Interestingly, the stricter the commuting flow, density or luminosity threshold used in each of the approaches (moving from left to right along the x-axes in Fig. 5(a)), the smaller is the estimated agglomeration wage premium. Moreover, and contrary to what Briant et al. (2010) find for France in the context of the Modifiable Areal Unit Problem (MAUP), the difference in results when using different approaches/thresholds to delineate metro areas is not much smaller than the difference in results arising from misspecification of the regression equation. On Java-Bali, the maximum absolute difference in the estimated agglomeration wage premium is an order of magnitude larger when using different approaches/thresholds to delineate metro areas (0.093) compared to the bias arising from not including any controls (maximally 0.047).²⁴ Compared to France, Java-Bali stands out in terms of its rapid urbanization, which is characterized by fast expanding metro areas. This may be an important explanation for why, in our case, results are at least as sensitive to the choice of approach and threshold(s) to define metro areas as to the choice of control variables to include.

5.3. Explaining the effect of the choice of approach and thresholds

To understand the observed patterns in our results, it is useful to consider that the estimated agglomeration wage premium, $\hat{\beta}$, that we obtain by estimating equation [1] using ordinary least squares (OLS) can be written as:

$$\hat{\beta} = \text{Cov}(\hat{S}, \hat{w}) / \text{Var}(\hat{S}) \quad (2)$$

where \hat{S} and \hat{w} are vectors of the (natural log) levels of experienced density and the nominal wage associated with each worker respectively, each “net of the effect of” the other controls included in equation [1].²⁵ Because we use the same fixed sample of urban workers in all regressions, as well as the exact same set of controls, the choice of approach and threshold(s) used to delineate metro areas only affects experienced density, \hat{S} , while leaving \hat{w} unaffected.

²⁴ Figure A2.3 in Appendix 2 reports results when including experienced density as the sole independent variable in the regression, i.e., excluding all control variables and industry fixed effects from equation [1].

²⁵ To be more specific, $\hat{S}(\hat{w})$ is a vector that stacks the residuals from a regression of the natural log of experienced density (natural log of the nominal wage) on the control variables, α_j, X_i, X_d , from equation [1]. For ease of exposition, we simply refer to density and nominal wages from now on in the main text, understanding these to mean density and nominal wages net of the effect of the included controls.

Now, for any given approach, the use of a lower (*de facto*) density threshold tends to result in the aggregation of new districts to the peripheries of existing metro areas and/or the emergence of new metro areas through the aggregation of districts that were not previously part of any metro at a higher threshold (see Section 4). By construction, the new districts that are added to the peripheries of existing metro areas tend to be less dense than the districts that were already part of the metro area. At the same time, they tend to be denser than the non-metro districts that remain. Similarly, the newly formed metro areas tend to involve the aggregation of one or more districts to a denser district, where all of the aggregated districts tend to be denser than the non-metro districts that remain. As a result, the use of a lower (*de facto*) density threshold tends to compress the distribution of experienced density across individual workers, by, in effect, averaging the density of the densest districts with the density of the less dense districts with which they form metro areas, while leaving the density of the least dense districts, which tend not to undergo aggregation into metro areas, unchanged. As a consequence, the variance of experienced density, $\text{Var}(\hat{S})$ in equation [2] tends to fall.

Moreover, in the presence of positive agglomeration effects, it is the higher paid nominal wage workers who are disproportionately present in the densest districts, and which, therefore, are most likely to see a reduction in their measured levels of experienced density as a result of the lower (*de facto*) density threshold. This implies that the use of a lower threshold also tends to reduce the positive covariance of experienced density with nominal wages across workers. Hence, $\text{Cov}(\hat{S}, \hat{w})$ in equation [2] tends to fall as well.²⁶ As such, for any given approach, the estimated agglomeration wage premium, $\hat{\beta}$, will increase (decrease) when using a lower (*de facto*) density threshold to define metro areas if this results in the variance of experienced density across workers declining faster (slower) in percentage terms than the covariance of experienced density and nominal wages across workers.

On Java-Bali, the above-outlined effect(s) of lowering the (*de facto*) density threshold are exactly borne out in the data. This can be seen in Fig. 6 which, for each approach to delineating metro areas, combines two scatterplots. The first is a scatterplot of the (natural log) of the nominal wages of individual workers against the natural log of experienced density when using a low (*de facto*) density threshold (the black markers). Meanwhile, the second is the same scatterplot but instead simply using each worker's experienced density calculated based on treating each administrative district as its own metro area (the gray markers), which corresponds to setting a very high (*de facto*) density threshold, in which there is no aggregation of districts. Relaxing the threshold associated with any given approach indeed results in a strong compression in the distribution of density across individual workers, which affects most severely those workers that are at the upper tail of the density distribution and which also, on average, have high nominal wages. As a result, both the variance of experienced density and the covariance of experienced density with nominal wages across workers fall. And, since the variance falls more than the covariance (see the numbers reported in the notes to Fig. 6), the fitted regression line, the gradient of which is equal to the estimated agglomeration wage premium, tilts upwards, becoming steeper. Moreover, this effect of lowering the (*de facto*) den-

²⁶ In particular: $\text{Cov}(\hat{S}, \hat{w}) = \frac{1}{N} \sum_{i=1}^N (\hat{S}_i - \mu_{\hat{S}})(\hat{w}_i - \mu_{\hat{w}})$, where \hat{S}_i and \hat{w}_i denote the natural logs of experienced density and the nominal wage respectively for worker i , each net of the effect of the control variables (see footnote 25), and $\mu_{\hat{S}}$ and $\mu_{\hat{w}}$ the respective means of these variables. Upon using a lower (*de facto*) density threshold, $(\hat{S}_i - \mu_{\hat{S}})$ falls most for workers with the highest values of $(\hat{w}_i - \mu_{\hat{w}})$, who are the workers employed in the densest districts with the highest nominal wages. It goes up for workers employed in less dense districts that are aggregated to more dense districts, but these tend to also have lower values of $(\hat{w}_i - \mu_{\hat{w}})$. Note that this covariance will, however, never switch sign: using a lower (*de facto*) density threshold to aggregate districts into metro areas will never result in the highest paid individuals being disproportionately present in the lowest density areas.

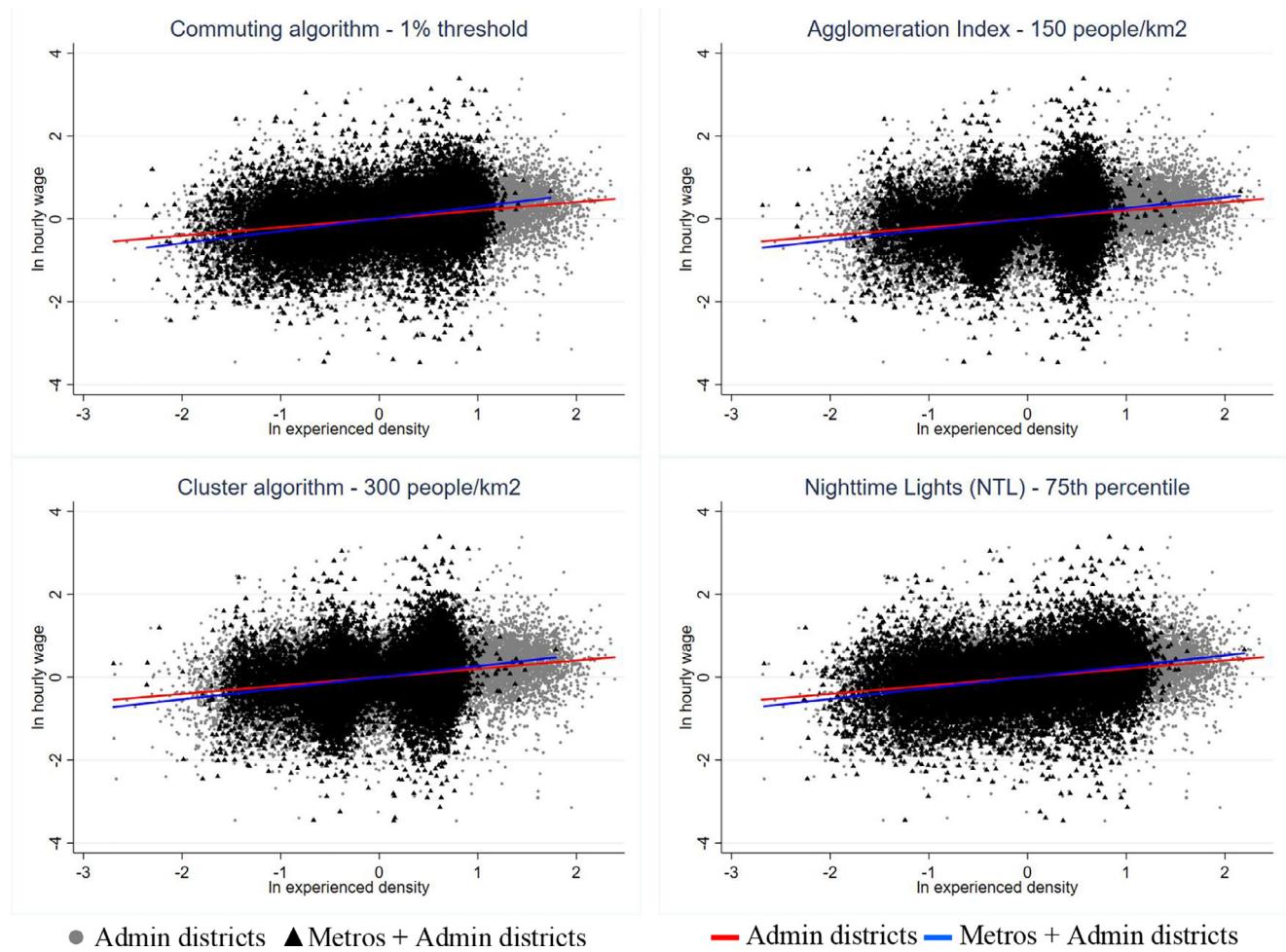


Fig. 6. Scatterplots of the nominal wage against experienced density for different approaches and thresholds – On Java-Bali.

Notes: The scatterplots are partial scatterplots which control for industry fixed effects, observable individual worker characteristics and island-group effects – i.e. they correspond to regression equation [1] and the results shown in Fig. 5(a) for the corresponding metro area definitions. As in that figure, for both the AI and cluster algorithms, the (core/overall) population threshold is kept fixed at 50,000, while the travel time threshold for the AI is also kept fixed at 60 min. The variance of $\ln(\text{experienced density})$ is 0.81 when each administrative district is treated as its own metro area without aggregation, while it is 0.49, 0.34, 0.35 and 0.56 when using the commuting, AI, cluster and NTL approaches respectively for the thresholds shown in this figure. The covariance of $\ln(\text{experienced density})$ and $\ln(\text{hourly wage})$ is 0.16 when each district is treated as its own metro area, while it is 0.14, 0.09, 0.10 and 0.15 when using the commuting, AI, cluster and NTL approaches respectively for the thresholds shown in this figure.

sity threshold used to define metro areas is stronger the greater is the reduction in the threshold. This explains why, in Fig. 5(a) for each approach, we find that the estimated agglomeration wage premium tends to keep increasing with the use of an ever more relaxed definition of metro areas.

By contrast, for the rest of Indonesia, most urban areas remain well-confined within administrative district boundaries, even when we (substantially) lower the (*de facto*) density threshold associated with each approach to defining metro areas. And, where the approaches do lead to the aggregation of districts into metro areas, these remain small (i.e. composed of an average of just over two districts) compared to those identified for Java-Bali (see Section 4). Moreover, the experienced density of the typical multi-district metro is often very similar to that of its densest district, where most workers are employed.²⁷ As a consequence,

regardless of the approach used, a comparison of the same two scatterplots as in Fig 6, one based on metro areas defined using a low (*de facto*) density threshold, and one based on simply treating each district as its own metro area, shows hardly any change in the distribution of experienced density across workers (Fig. 7). As a result, neither the variance of (natural log) experienced density nor the covariance of this density with the (natural log) nominal wage is much affected by using a lower (*de facto*) density threshold (see the reported variance and covariance in the notes to Fig. 7). Together with the much weaker association between nominal wages and density that we found off Java-Bali in the first place, this explains why we see hardly any effect of the choice of threshold,

²⁷ This is in no small part due to experienced density being better suited than naive density to capturing how close the typical person is to other people when the population is unevenly distributed (see also Duranton and Puga, 2020, p 3). A corollary of this is that it is relatively robust to the drawing of boundaries that lead to the definition of metro areas that are “too large” in the sense that they

start to include large and relatively sparsely populated rural areas. In such cases, mostly found off Java-Bali in Indonesia, experienced density hardly changes for the many workers living in the densest part of the now extended metro areas, only increasing it for the (relatively) few workers in the newly added peripheral districts of the metro.

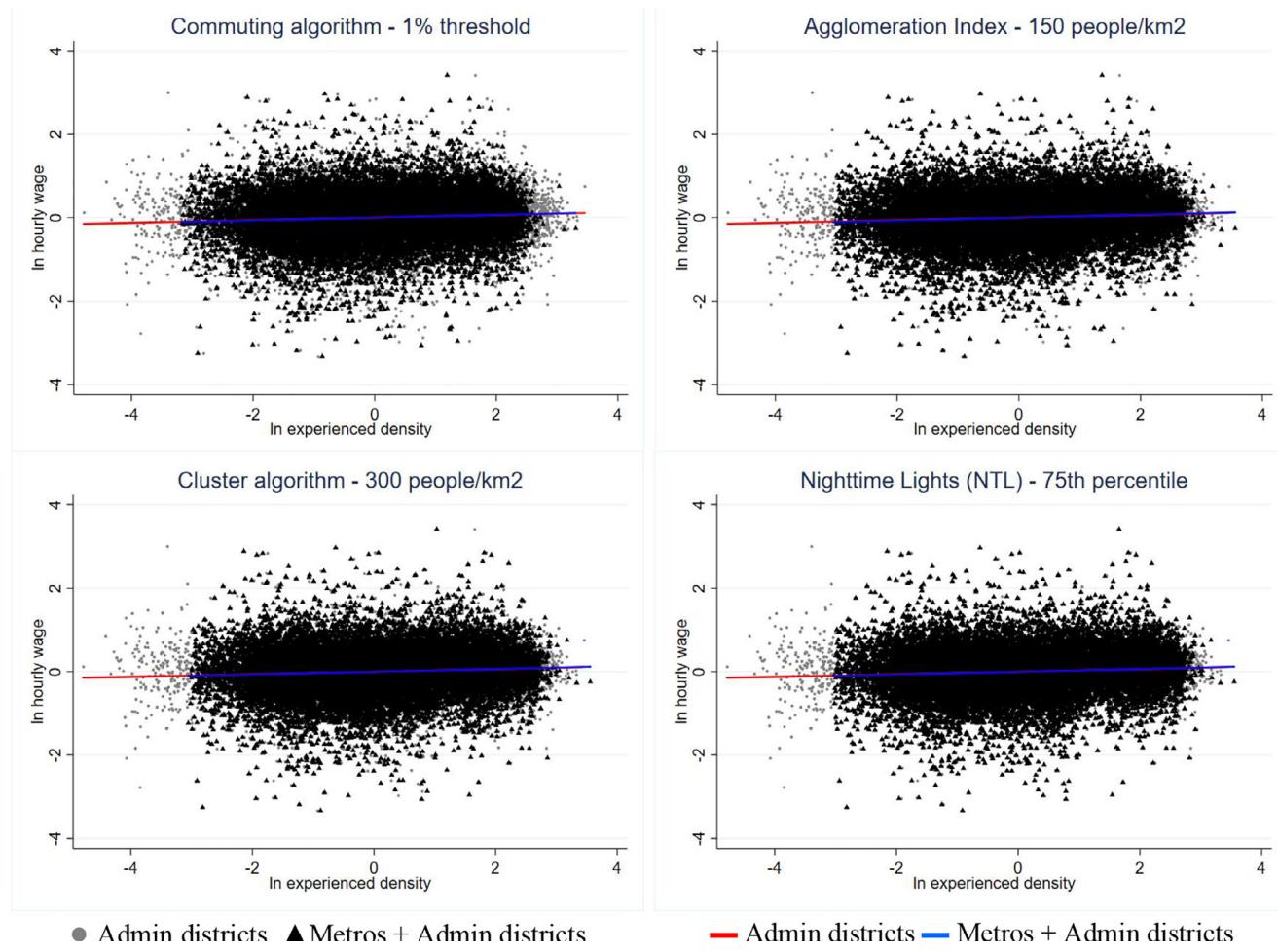


Fig. 7. Scatterplots of the nominal wage against experienced density for different approaches and thresholds – Off Java-Bali.

Notes: The scatterplots are partial scatterplots which control for industry fixed effects, observable individual worker characteristics and island-group effects – i.e. they correspond to regression equation [1] and the results shown in Fig. 5(b) for the corresponding metro area definitions. As in that figure, for both the AI and cluster algorithms, the (core/overall) population threshold is kept fixed at 50,000, while the travel time threshold for the AI is also kept fixed at 60 min. The variance of $\ln(\text{experienced density})$ is 1.98 when each administrative district is treated as its own metro area without aggregation, while it is 1.77, 1.98, 1.95 and 2.00 when using the commuting, AI, cluster and NTL approaches respectively for the thresholds shown in this figure. The covariance of $\ln(\text{experienced density})$ and $\ln(\text{hourly wage})$ is 0.06 when each district is treated as its own metro area, while it is 0.05, 0.06, 0.06 and 0.06 when using the commuting, AI, cluster and NTL approaches respectively for the thresholds shown in this figure.

or, indeed, approach used to delineate metro areas, on the estimated agglomeration wage premium outside of Java-Bali.²⁸

As a caveat to the above (mechanical) explanation for the observed patterns in the size of the estimated agglomeration wage premium, it should be noted that the choice of approach and threshold(s) may have other consequences for the estimation of the premium that are not obvious in our results. A shortcoming of our approach to estimating the wage premium which applies across all four metro area delineation approaches, is that we cannot fully exclude the possibility of not having adequately controlled for all worker, firm, or even location-specific characteristics that affect an individual worker's nominal wage and that are

also disproportionately present in the denser parts of Indonesia. And, importantly from the perspective of the results discussed in this section, the extent of the resultant endogeneity problem(s) could conceivably be scale dependent.

6. A tentative exploration of the factors behind Indonesia's agglomeration wage premium

In this final section, we tentatively explore possible explanations for the significant estimated agglomeration wage premium and probe why this is so much larger for Java-Bali than for the rest of Indonesia. To keep the discussion focused, we only show results where experienced density and all other agglomeration related variables are constructed using Indonesia's administrative district boundaries. This is because, as noted in Section 5.2, the large difference in the estimated wage premium between Java-Bali and the rest of Indonesia exists irrespective of how we define the spatial units. All insights discussed in this section also hold when instead using one of the four approaches to delineating metro areas.

²⁸ A further factor that contributes to the difference in results on and off Java-Bali is that districts, which provide the “building blocks” for constructing metro areas using the four approaches, are, on average, smaller on Java-Bali. This makes it more likely off Java-Bali that urban areas are contained within district boundaries and that delineated spatial units contain possibly large amounts of rural area, particularly as districts are aggregated. As indicated in footnote 27, experienced density is robust to the delineation of spatial units that are “too large” in this sense.

Our tentative exploration consists, firstly, of adding several explanatory variables to our baseline regressions (see equation [1]) for both Java-Bali and the rest of Indonesia, each of which captures a different potential reason for the existence of an agglomeration wage premium. These variables are measures of both domestic and international market access; measures of both a district's level of specialization in the industry in which the worker is employed and a district's overall level of industrial diversity; the share of manufacturing in district employment and the average level of total factor productivity (TFP) of a district's manufacturing plants; and, finally, the share of skilled workers in a district's total employment, where we define a skilled worker as one who has completed at least general senior high school. Should a variable be important in explaining the difference in the estimated premiums between Java-Bali and the rest of Indonesia, then controlling for it should both reduce the estimates of the premium and lead to their convergence.²⁹

Second, we probe whether the existence of heterogeneous effects of density on workers' productivity may help to explain the difference in the estimated premium between Java-Bali and the rest of Indonesia. We do this by adding interactions of experienced density with dummy variables for whether or not a worker is skilled, whether or not a worker is employed in the manufacturing sector, and whether or not a worker is employed in a district that has a level of experienced density in excess of the median for all Indonesian districts. This last variable allows us to capture the possibility of a non-linear effect of density on productivity (see also Kline and Moretti, 2014).

Table 1 presents the results from this exercise. In columns [1] and [2] we see that the addition of the new explanatory variables lowers the estimated agglomeration wage premium both on and off Java-Bali. For Java-Bali, the estimated premium falls by almost 30% from 0.202 in our baseline regression to 0.142, while for the rest of Indonesia it falls from 0.032 to -0.001, becoming statistically insignificant in the process. The results also suggest some interesting differences in the factors behind the agglomeration wage premium on and off Java-Bali. A larger employment share in the manufacturing sector, higher average local manufacturing TFP, as well as a more diverse overall industrial structure are all significantly positively associated with higher nominal wages for both parts of Indonesia. These effects are suggestive of the existence of positive spillovers from a district's manufacturing sector to the rest of its economy, as well as of the existence of Jacob's style urbanization economies (Jacobs, 1969). But, on Java-Bali, the association is stronger in the cases of both industrial diversity and manufacturing TFP, while the association with the share of manufacturing workers is stronger in the rest of Indonesia. Moreover, only on Java-Bali do we find some evidence of a significant positive association between (international) market access, which we proxy by travel time to the nearest of 35 international ports, and nominal wages. By contrast, only in the rest of Indonesia do we find the share of skilled workers in employment to be significantly positively related to higher local nominal wages, which may be suggestive of the presence of human capital externalities (Rauch, 1993; Moretti, 2004; Dingel et al., 2020).

When also including the interaction terms with experienced density, the results in columns [3] and [4] show that there are significant positive interaction effects of experienced density with both a worker's skill level and whether or not a worker is employed in a district of above

²⁹ We firstly choose these variables because they capture some of the most prominent potential reasons for the existence of the productivity benefits associated with density (see, for example, Rosenthal and Strange, 2004; Duranton and Puga, 2020). And, second, because districts on Java-Bali are markedly different to districts in the rest of Indonesia when it comes to most of these variables. In particular, districts on Java-Bali have substantially better market access, a substantially higher share of manufacturing employment, much more productive manufacturing plants, and are home to a much more diverse set of industries. Full details on the methods of construction for each of these variables beyond those contained in the notes to Table 1 are available on request.

median density, but not with whether a worker is employed in manufacturing.³⁰ The effects for skill level and above median density are present for both Java-Bali and the rest of Indonesia but are stronger for the former. For skilled workers on Java-Bali, a doubling of density is associated with a 17.6% increase in the nominal wage, which is 8.4 percentage points higher than the 9.2% wage increase that we find for unskilled workers. By contrast, skilled workers in the rest of Indonesia are only estimated to receive a two percentage points higher wage increase than unskilled workers in response to a doubling of density.³¹ In the meantime, a worker is estimated to receive a nominal wage increase of 9.2% as a result of a doubling of density on Java-Bali, where, in fact, all districts have levels of experienced density in excess of the median for all Indonesian districts.³² For the rest of Indonesia, where only 33% of districts are of above median density, a doubling of density is instead associated with a nominal wage increase that is 6.3 percentage points higher for workers employed in districts of above median density than it is for workers employed in districts of below median density, for which the effect of density on wages is insignificant.

The above results suggest the possible importance of two explanations for the large difference in the estimated agglomeration wage premium between Java-Bali and the rest of Indonesia. The first is that skilled workers, who tend to be disproportionately present in denser areas both on and off Java-Bali, benefit more from density on Java-Bali than they do off Java-Bali. Second, we find that density appears to be non-linearly associated with wages, and, therefore, productivity, in Indonesia. Hence, we only find evidence of a significant agglomeration wage premium in districts of above median density. These districts are far more prevalent on Java-Bali (see also Appendix 2, Fig. A2.2). It might simply be the case that outside of Java-Bali, the scope for agglomeration economies is still, given the much lower levels of density, limited.

While the above results are suggestive, it is important to emphasize that they do not provide the final word on the sources of Indonesia's agglomeration wage premium and its estimated difference between Java-Bali and the rest of the country. In particular, we cannot exclude the possibility that our results are driven by unobserved worker-, firm-, or place-based characteristics that systematically differ between the more and less dense places in Indonesia (see also the discussion at the end of Section 5.3). We also do not observe an individual worker's employer, which leaves us unable to analyze the potential productivity benefits of denser places associated with the better matching of the most productive workers with the most productive firms (see, for example, Behrens et al., 2014). Furthermore, the variables that we have included above to capture some of the potential sources of the agglomeration wage premium are proxies and, as such, may not fully capture these sources. We, unfortunately, lack convincing sources of exogenous variation to fully address these issues, and pin down how each of the different variables that we have introduced causally affects wages. As such, our results are best viewed as a preliminary step, intended to fuel future research, towards uncovering the sources of Indonesia's agglomeration wage pre-

³⁰ Although the estimated coefficient on the interaction of (natural log) experienced density with whether a worker is employed in manufacturing is significantly negative in column [4] for the rest of Indonesia, it is only so at the 10% level. Furthermore, this finding of a significant effect is not robust to small changes in the regression specification – for example, dropping one or more of the additionally included explanatory variables.

³¹ This effect is, nevertheless, still sizable given the complete absence of any evidence of an effect of density on unskilled workers' wages outside of Java-Bali.

³² This implies that the estimated elasticity of nominal wages with respect to density for above median density districts in column [3] of Table 1 is simply given by the estimated coefficient on effective density – i.e. there is no need to interact effective density with the dummy for whether or not a worker is employed in a district of above median density.

Table 1
Potential sources of the agglomeration wage premium, on and off Java-Bali

VARIABLE	[1] Java-Bali	[2] off Java-Bali	[3] Java-Bali	[4] off Java-Bali
ln (experienced density)	0.142*** (0.036)	-0.001 (0.013)	0.092** (0.036)	-0.030 (0.031)
× manufacturing worker (=1 if yes)	- -	- (0.021)	-0.013 (0.018)	-0.031* (0.010)
× skilled worker (=1 if yes)	- -	- (0.014)	0.084*** (0.014)	0.020** (0.010)
× district denser than median (=1 if yes)	- -	- (0.036)	- (0.036)	0.063* (0.036)
ln (domestic market access)	-0.012 (0.010)	-0.012 (0.013)	-0.010 (0.009)	-0.013 (0.013)
ln (travel time to nearest port)	-0.058* (0.030)	-0.006 (0.013)	-0.055* (0.030)	-0.006 (0.013)
% workers in same sector	0.001 (0.002)	-0.001 (0.002)	0.000 (0.002)	-0.001 (0.002)
Industrial diversity	0.022*** (0.005)	0.005* (0.003)	0.022*** (0.004)	0.004 (0.003)
% workers in manufacturing	0.005* (0.003)	0.014*** (0.005)	0.006** (0.003)	0.017*** (0.006)
(mean) ln manufacturing TFP	0.107*** (0.029)	0.039* (0.018)	0.108*** (0.029)	0.032* (0.018)
% skilled workers	-0.002 (0.002)	0.003** (0.002)	-0.002 (0.002)	0.003* (0.002)
Adjusted R ²	0.406	0.287	0.409	0.290
N	17,604	12,686	17,604	12,686

Notes: ***, **, * denote significance at the 1%, 5%, and 10% levels respectively. Standard errors clustered at the district level in parentheses. Domestic market access is measured as the travel time discounted sum of the populations of all Indonesian districts, excluding the population of the district for which it is calculated. As with the calculation of the travel time to the nearest international port, travel times are estimated based on digitized maps of Indonesia's road and ferry networks (for further details of the calculation of the domestic and international market access variables see Roberts, Gil Sander and Tiwari, eds., 2019, Annex 3A). % workers in same sector refers to workers employed in the same of nine main sectors defined in the 1990 Indonesian Standard Industrial Classification. Industrial diversity is calculated as -100 × HHI, where HHI denotes the Herfindahl-Hirschman Index of sectoral employment shares based on two-digit sectors in the 2009 Indonesian Standard Industrial Classification. TFP denotes total factor productivity, the calculation of which is detailed in [World Bank \(2018\)](#). All regressions also include industry fixed effects, island-group fixed effects and individual worker characteristics as controls following the specification of equation [1]. In column [3], no estimated coefficient is reported for the interaction of ln (experienced density) with the dummy for districts whose experienced density exceeds that of the median Indonesian district. This is because all districts on Java-Bali have levels of experienced density in excess of the median.

mium and the reasons for its difference between Java-Bali and the rest of the country.

7. Conclusion

A variety of approaches to delineating metropolitan areas have been developed. This paper highlights the importance of the choice of approach and associated threshold(s) used to delineate metropolitan areas. Our results show that this choice can have non-trivial consequences, not only in terms of the basic descriptions that they provide of the metropolitan area landscape (notably the number of metro areas, their size, and the shares of urban population that they cover), but also when including variables defined at the metropolitan area level as independent variables in a regression context.

We focused our investigation on the particular case of Indonesia, a rapidly urbanizing emerging economy with fast expanding cities that are sprawling across official administrative boundaries. Moreover, Indonesia's distinct geography with its split between densely populated Java-Bali and the much less densely populated rest of the country allowed us to perform our comparisons both in a setting of (very) high average population density, and a setting of much lower average population density. Our findings point to two important conclusions.

First, in settings where commuting flow data is absent, great care is required in using non-commuting data approaches to generate proxy metropolitan areas. This is because there is no threshold or set of thresholds that allow such approaches to generate descriptions of Indonesia's metro area landscape that resemble those obtained using the commuting algorithm. In the case of Java-Bali, the non-commuting based approaches easily generate implausibly large metro areas, especially when using lower density or luminosity thresholds. Meanwhile, for the rest of Indonesia, these same approaches instead identify far fewer metro areas than would be suggested to exist based on actual commuting flows between districts.

Second, and finally, when estimating the agglomeration wage premium using a density measure, researchers should ideally report results using different density measures, each calculated based on a different definition of metro areas. This is, in our view, particularly salient in a context of rapid urbanization, like the one found on Java-Bali, where cities have been growing rapidly, both in terms of overall population as well as, importantly, spatial extent. Administrative boundaries in such a case are likely to no longer accurately reflect the extents of metro areas, which we have shown can have non-trivial implications for the estimation of the agglomeration wage premium.

Appendix 1. Additional descriptive results

Table A1.1
Comparison of key statistics – selected results

Threshold	No. metros (Java-Bali)	No. metro districts		Population		Largest metro Name (no. districts)	Threshold	No. met- ros Bali	No. metro districts		Population		Largest metro Name (no. districts)			
		Total	Mean	Total (mil.)	% IDN				Total	Mean	Total (mil.)	% IDN				
Commuting algorithm (% commuting)																
27.0%	1 (1)	2	2.0	3.05	1.2	Bandung (2)	1500-500	38	4.2	62.53	24.8	Jakarta (15)				
23.0%	2 (2)	4	2.0	6.71	2.7	Jakarta Selatan (2)	1500-100	9 (8)	38	4.2	62.53	24.8	Jakarta (15)			
21.0%	3 (2)	6	2.0	10.88	4.3	Medan (2)	1500-50	9 (8)	38	4.2	62.53	24.8	Jakarta (15)			
17.5%	4 (3)	8	2.0	12.36	4.9	Medan (2)	1500-5	9 (8)	39	4.3	63.81	25.3	Jakarta (15)			
15.5%	8 (5)	17	2.1	19.31	7.7	Jakarta Selatan (3)	1100-500	12 (10)	51	4.3	72.51	28.8	Jakarta (16)			
12.0%	14 (8)	31	2.2	32.81	13.0	Surabaya (3)	1100-100	12 (10)	51	4.3	72.51	28.8	Jakarta (16)			
10.0%	24 (17)	58	2.4	64.51	25.6	Bogor (2)	1100-5	12 (10)	51	4.3	72.51	28.8	Jakarta (16)			
9.5%	26 (16)	70	2.7	75.54	30.0	Jakarta (11)	700-500	13 (11)	74	5.7	96.00	38.1	Jakarta (20)			
8.0%	32 (17)	86	2.7	88.18	35.0	Jakarta (11)	700-100	13 (11)	74	5.7	96.00	38.1	Jakarta (20)			
7.0%	39 (18)	103	2.6	97.09	38.5	Jakarta (13)	700-5	13 (11)	74	5.7	96.00	38.1	Jakarta (20)			
6.0%	42 (17)	113	2.7	100.72	40.0	Jakarta (13)	300-500	15 (8)	129	8.6	143.26	56.8	Central-East Java (55)			
4.0%	52 (19)	141	2.7	115.60	45.9	Jakarta (13)	300-100	17 (8)	133	7.8	144.07	57.2	Central-East Java (55)			
2.0%	68 (21)	192	2.8	140.27	55.7	Jakarta (15)	300-5	18 (8)	135	7.5	144.24	57.2	Central-East Java (55)			
1.0%	80 (21)	245	3.1	169.20	67.1	Jakarta (21)	Cluster algorithm (tax density – population)									
0.5%	82 (15)	294	3.6	193.02	76.6	Jakarta (21)	1500-500	5 (4)	25	5.0	49.01	19.4	Jakarta (12)			
Nighttime lights thresholding (percentile of national range)																
95 th	3 (3)	13	4.3	26.18	10.4	Jakarta (9)	1500-100	6 (5)	28	4.7	53.28	21.1	Jakarta (13)			
90 th	5 (4)	21	4.2	38.16	15.1	Jakarta (11)	1500-50	6 (5)	28	4.7	53.28	21.1	Jakarta (13)			
80 th	8 (6)	41	5.1	64.42	25.6	Jakarta (15)	300-500	8 (6)	48	6.0	71.99	28.6	Jakarta (16)			
60 th	9 (6)	69	7.7	92.05	36.5	West Java (26)	300-100	12 (8)	90	7.5	112.27	44.5	W-Cent Java (29)			
40 th	9 (5)	89	9.9	111.68	44.3	W-Cent Java (31)	300-50	12 (8)	99	8.3	121.41	48.2	W-Cent Java (29)			
25 th	9 (4)	109	12.1	125.66	49.9	W-Cent Java (33)	150-500	9 (6)	55	6.1	76.14	30.2	W-Cent Java (21)			
20 th	9 (4)	114	12.7	129.12	51.2	W-Cent Java (33)	150-100	11 (4)	105	9.5	123.28	48.9	W-Cent Java (37)			
5 th	8 (3)	119	14.9	132.13	52.4	Java (92)	150-50	13 (5)	123	9.5	139.11	55.2	W-Cent Java (38)			
Agglomeration Index (density - population)																

Source: (Urban) population data from Indonesia's 2014 National Socio-Economic Survey (*Survei Sosial Ekonomi Nasional*, SUSENAS).

Note: IDN = Indonesia. mil. = Million. W-Cent Java = West-Central Java. For the Agglomeration Index, the travel time threshold is held fixed at 60 min. Density refers to population density (people per square km). Population refers to overall population for the cluster algorithm and the core population for the Agglomeration Index.

Appendix 2. Additional results

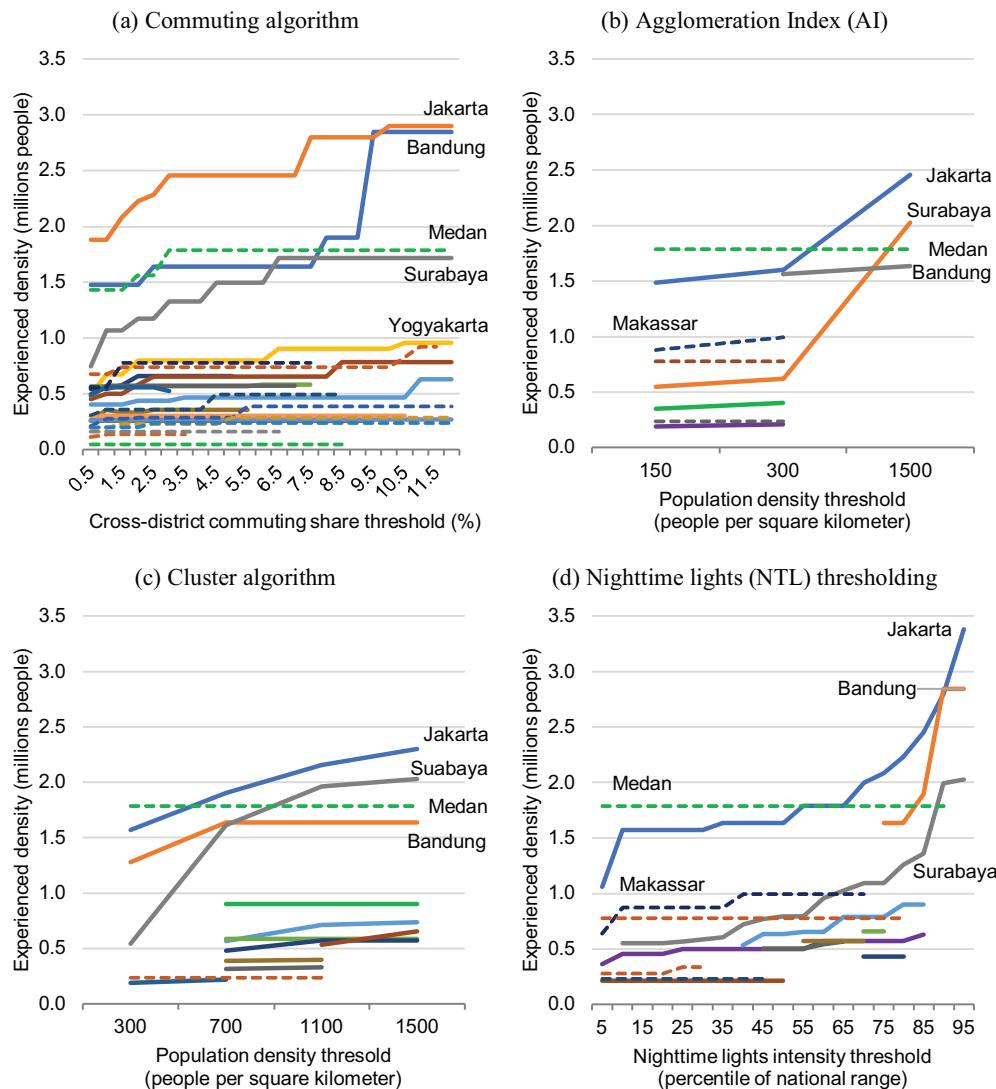


Fig. A2.1. Changes in density of identified metro areas across thresholds by approach. Source: Experienced density calculated using *Landsat-2012* gridded population data. Notes: Each depicted line shows how the density of an identified metro area, as measured by the average number of people within 10 km of the average person (De La Roca and Puga, 2017), varies with each approach's most important threshold. In addition to the thresholds shown on the x-axis, we used a minimum core population threshold of 50,000 and a travel time threshold of 60 min in panel (b) and a minimum overall population threshold of 50,000 in panel (c). Solid and dashed lines indicate metro areas on and off Java-Bali, respectively.

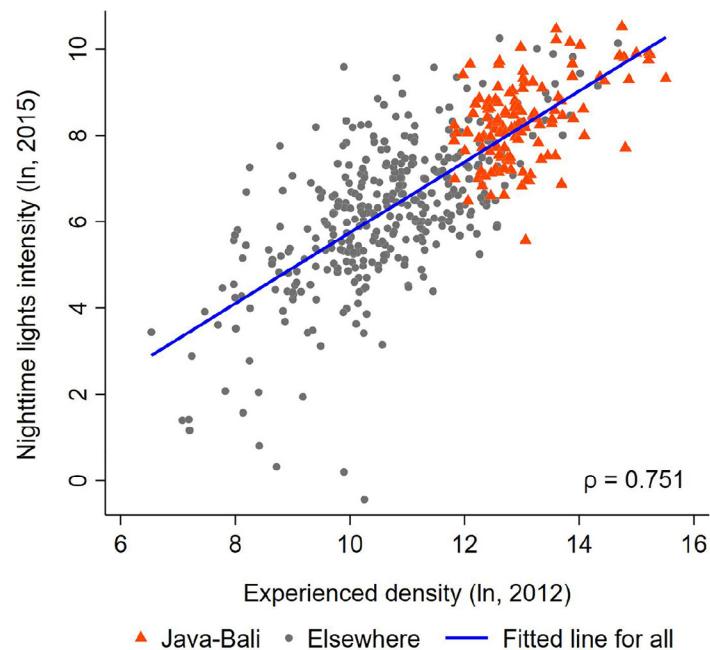
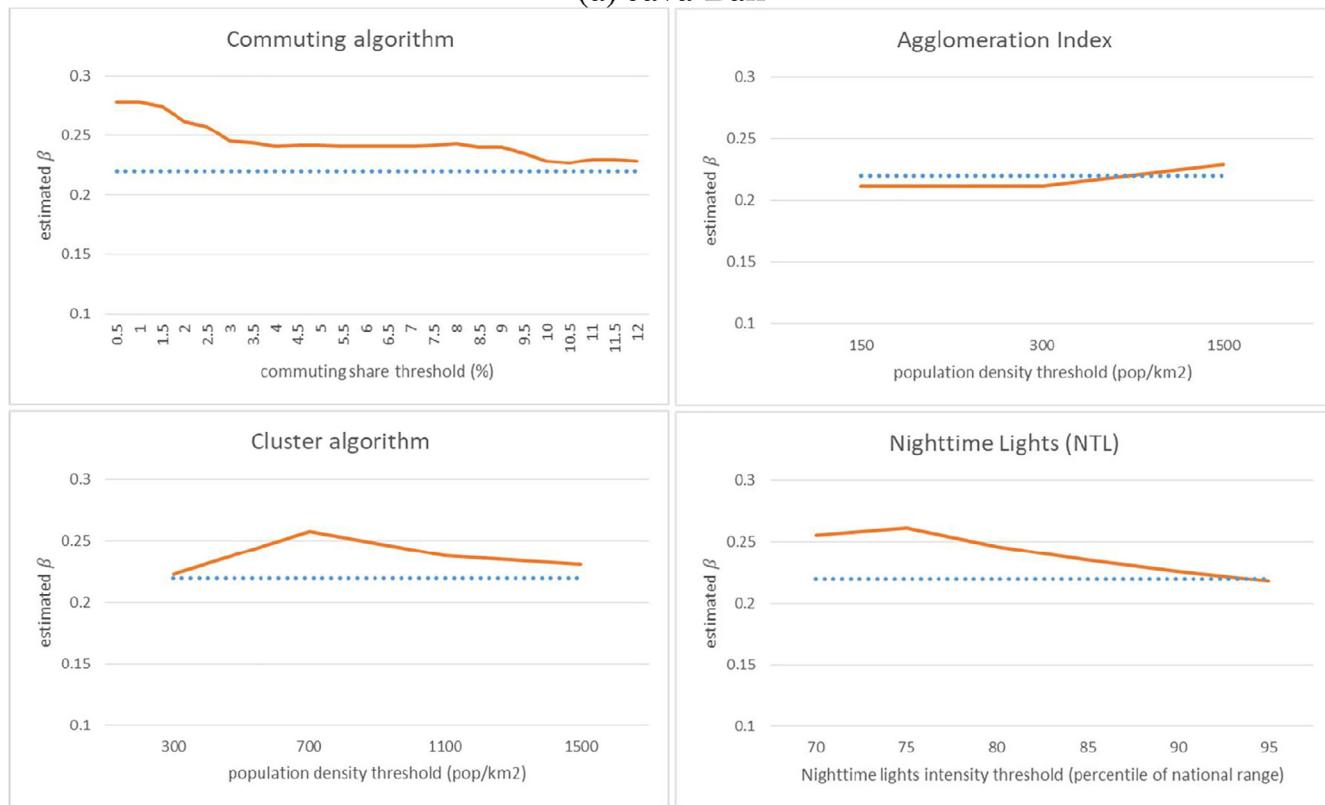


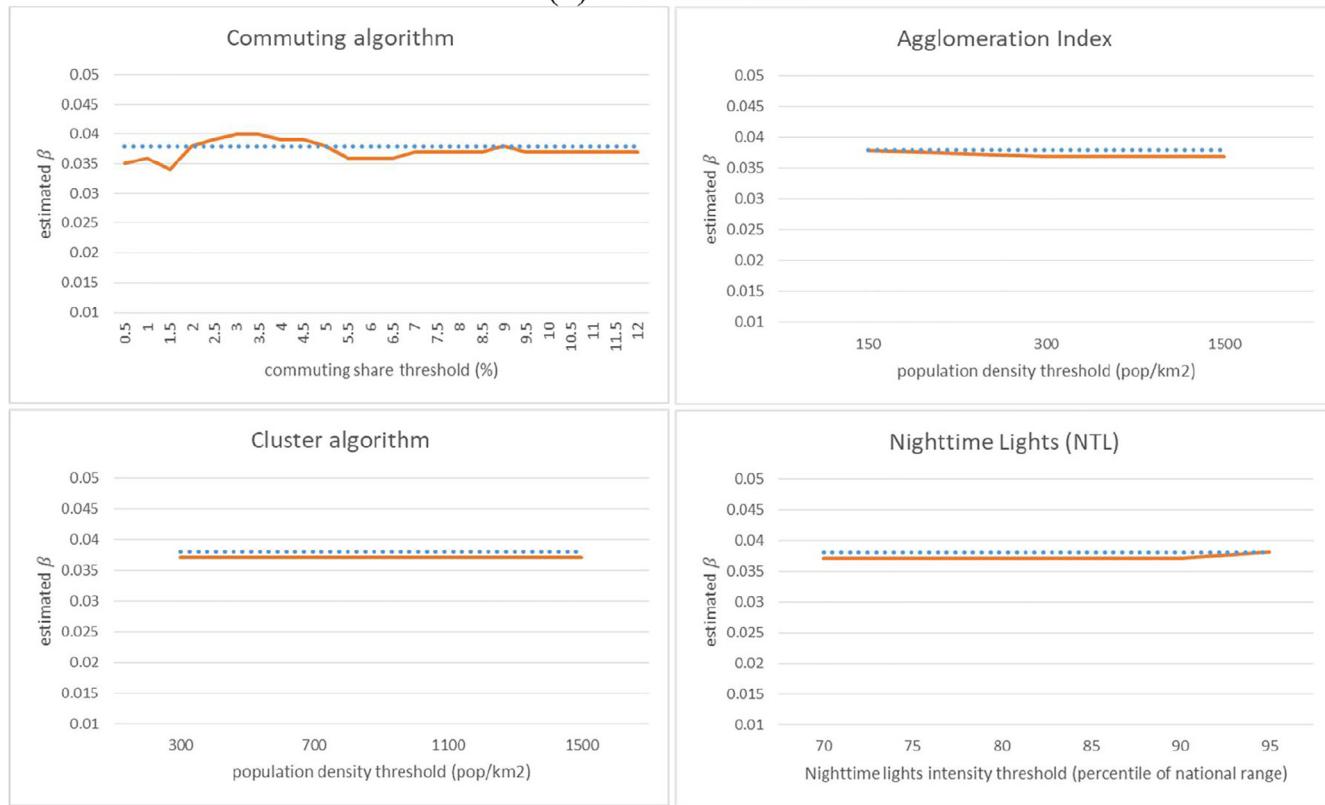
Fig. A2.2. Correlation between density and NTL luminosity – Indonesian districts.

Notes: Following De La Roca and Puga (2017), experienced density is defined as the average number of people within 10 km of the average person living in the district; NTL intensity is measured as the sum of the luminosity (DN values) across all pixels that are part of the district.

(a) Java-Bali



(b) Off Java-Bali

**Fig. A2.3.** Without any control variables included.

Notes: All the depicted estimates of the agglomeration wage premium are significant at the 5% level. Adding confidence bands to the graphs is possible but would make the patterns in the estimates less visible due to the resulting extended y-axis needed to accommodate the bands. For both the AI and cluster algorithms, the (core/overall) population threshold is kept fixed at 50,000; the travel time threshold for the AI is also kept fixed at 60 min.

CRediT authorship contribution statement

Maarten Bosker: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision. **Jane Park:** Software, Formal analysis, Investigation, Data curation, Visualization. **Mark Roberts:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition.

References

- Abrahams, A., Oram, C., Lozano-Gracia, N., 2018. Deblurring DMSP nighttime lights: a new method using gaussian filters and frequencies of illumination. *Remote Sens. Environ.* 210, 242–258.
- Ahfeldt, G.M., Pietrofani, E., 2019. The economic effects of density: a synthesis. *J. Urban Econ.* 111, 93–107.
- Arribas-Bel, D., García-López, M.A., Viladecans-Marsal, E., 2020. Building(s and) Cities: delineating urban areas with a machine learning algorithm. *J. Urban Econ.*
- Baragwanath, K., Goldblatt, R., Hanson, G., Khandelwal, A., 2020. Detecting urban markets with satellite imagery: an application to India. *J. Urban Econ.*
- de Bellefon, M.P., Combes, P.-P., Duranton, G., Gobillon, L., Gorin, C., 2020. Delineating urban areas using building density. *J. Urban Econ.*
- Behrens, K., Duranton, G., Robert-Nicoud, F., 2014. Productive cities: sorting, selection, and agglomeration. *J. Polit. Econ.* 122 (3), 507–553.
- Berg, C.N., Blankepoor, B., Li, Selod, H., 2017. Global Travel Time to Major Cities, Circa 2010. The World Bank, Washington, D.C. Unpublished manuscript under preparation.
- Briant, A., Combes, P., Lafourcade, M., 2010. Dots to boxes: do the size and shape of spatial units jeopardize economic geography estimations? *J. Urban Econ.* 67, 287–302.
- Center for International Earth Science Information Network (CIESIN); Columbia University, 2017. Documentation for the Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Settlement Points, Revision 01. Palisades. NASA Socioeconomic Data and Applications Center (SEDAC), NY. <http://dx.doi.org/10.7927/H4GX48H5>.
- Ch, R., Martin, D.A., Vargas, J.F., 2020. Measuring the size and growth of cities using nighttime light. *J. Urban Econ.*
- Chauvin, J.P., Glaeser, E., Ma, Y., Tobio, K., 2017. What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *J. Urban Econ.* 98 (C), 17–49.
- Ciccone, A., Hall, R., 1996. Productivity and the density of economic activity. *Am. Econ. Rev.* 86 (1), 54–70.
- Combes, P.-P., Gobillon, L., 2015. The empirics of agglomeration economies. In: Duranton, G., Henderson, J.V., Strange, W. (Eds.), *Handbook of Regional and Urban Economics*, Volume 5. Elsevier, Amsterdam, pp. 247–348.
- Combes, P.-P., Démurger, S., Li, S., 2015. Migration externalities in Chinese cities. *Eur. Econ. Rev.* 76, 152–167.
- Combes, P.-P., Démurger, S., Li, S., Wang, J., 2019. Unequal migration and urbanisation gains in China. *J. Dev. Econ.* 142, 102328.
- Danko, D.M., 1992. The digital chart of the world project. *Photogram. Eng. Remote Sens.* 58, 1125–1128.
- De La Roca, J., Puga, D., 2017. Learning by working in big cities. *Rev. Econ. Stud.* 84, 106–142.
- Dijkstra, L., Poelman, H., 2014. A Harmonised Definition of Cities and Rural Areas: The New Degree of Urbanization. Directorate-General for Regional and Urban Policy, European Commission, Brussels Regional Working Paper.
- Dijkstra, L., Florczyk, A.J., Freire, S., Kemper, T., Pesaresi, M., Schiavina, M., 2019. “Applying the Degree of Urbanization to the Globe: A New Harmonized Definition Reveals a Different Picture of Global Urbanization. European Commission, Directorate-General for Regional and Urban Policy Unpublished manuscript.
- Dingel, J.I., Misco, A., Davis, D.R., 2020. Cities, lights, and skills in developing economies. *J. Urban Econ.*
- Doll, C.N.H., 2008. CIESIN Thematic Guide to Night-Time Light Remote Sensing and Its Applications. Center for International Earth Science Information Network of Columbia University, Palisades, NY.
- Duranton, G., 2015a. A proposal to delineate metropolitan areas in Colombia. *Desarrollo y Sociedad* 15, 223–264.
- Duranton, G., 2015b. Growing through cities in developing countries. *World Bank Res. Observ.* 30 (1), 39–73.
- Duranton, G., 2016. Agglomeration effects in Colombia. *J. Reg. Sci.* 56 (2), 210–238.
- Duranton, G., Puga, D., 2020. The Economics of Density. *J. Econ. Perspect.* 34 (3) (forthcoming).
- Ellis, P., Roberts, M., 2016. Leveraging Urbanization in South Asia: Managing Spatial Transformation for Prosperity and Livability. World BankWorld Bank, Washington, DC.
- Elvidge, C.D., Kihn, K.E., Davis, E.R., 1996. Mapping city lights with nighttime data from the DMSP-OLS operational linescan system. *Photogramm. Eng. Remote Sens.* 63, 727–734.
- Galdo, V., Li, Y., Rama, M., 2020. Identifying urban areas by combining human judgement and machine learning: An application to India. *J. Urban Econ.*
- Glaeser, E.L., Henderson, J.V., 2017. Urban economics for the developing world: an introduction. *J. Urban Econ.* 98, 1–5.
- Henderson, J.V., Nigmatulina, D., Kriticos, S., 2020. Measuring urban economic density. *J. Urban Econ.*
- Holmes, T., Lee, S., 2010. Cities as six-by-six-mile squares: Zipf's law? In: Glaeser, E. (Ed.), *Agglomeration Economics*. The University of Chicago press, Chicago, IL, pp. 105–131.
- Imhoff, M.L., Lawrence, W.T., Stutzer, D.C., Elvidge, C.D., 1997. A Technique for using composite Dmsp/Ols “city lights” satellite data to map urban area. *Remote Sens. Environ.* 61 (3), 361–370.
- Jacobs, J., 1969. *The Economy of Cities*. Random House, New York.
- Kline, P., Moretti, E., 2014. Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority. *Quart. J. Econ.* 129 (1), 275–331.
- Marshall, A., 1890. *Principles of Economics*. Macmillan, London.
- Moreno-Monroy, A.I., Schiavina, M., Veneri, P., 2020. Metropolitan areas in the world. Delineation and population trends. *J. Urban Econ.*
- Moretti, E., 2004. Human capital externalities in cities. In: Henderson, J.V., Thisse, J.-F. (Eds.), *Handbook of Urban and Regional Economics*, Volume 4. North Holland, New York, pp. 2243–2291.
- OECD (Organisation for Economic Co-operation and Development), 2012. Redefining “Urban”: A New Way to Measure Metropolitan Areas. OECD Publishing.
- Office of Management and Budget, 2010. 2010 standards for delineating metropolitan and micropolitan statistical areas. *Fed. Regist.* 75, 37246–37252.
- Overman, H.G., Venables, A.J., 2005. Cities in the developing world. London School of Economics and Political Science, London Center for Economic Performance Discussion Paper 695.
- Pinkovskiy, M.L., 2013. Economic Discontinuities at Borders: Evidence from Satellite Data on Lights at Night. Massachusetts Institute of Technology, Cambridge, MA Unpublished.
- Quintero, I., Roberts, M., 2018. Explaining Spatial Variations in Productivity: Evidence from 16 Latin American and Caribbean Countries. World Bank, Washington, DC Policy Research Working Paper 8560.
- Rauch, J.E., 1993. Productivity gains from geographic concentration of human capital: evidence from the cities. *J. Urban Econ.* 34 (3), 380–400.
- Roberts, M., 2018a. “The many dimensions of urbanization and the productivity of cities in Latin America and the Caribbean. In: Ferreyra, M.M., Roberts, M. (Eds.), *Raising the Bar for Productive Cities in Latin America and the Caribbean*. The World Bank, Washington, D.C., pp. 49–85.
- Roberts, M., 2018b. The empirical determinants of city productivity. In: Ferreyra, M.M., Roberts, M. (Eds.), *Raising the Bar for Productive Cities in Latin America and the Caribbean*. The World Bank, Washington, D.C., pp. 89–115.
- Roberts, M., Gil Sander, F., Tiwari, S., 2019. Time to ACT: Realizing Indonesia’s Urban Potential. World Bank, Washington, DC.
- Rosenthal, S.S., Strange, W.C., 2004. Evidence on the nature and sources of agglomeration economies. In: Henderson, J.V., Thisse, J.-F. (Eds.), *Handbook of Urban and Regional Economics*, Volume 4. North Holland, New York, pp. 2119–2171.
- Rozenfeld, H.D., Rybski, D., Gabaix, X., Makse, H.A., 2011. The area and population of cities: new insights from a different perspective on cities. *Am. Econ. Rev.* 101, 2205–2225.
- Schiavina, M., Freire, S., MacManus, K., 2019. GHS Population Grid Multitemporal (1975, 1990, 2000, 2015) R2019A. European Commission, Joint Research Centre (JRC).
- Small, C., 2019. Multisensor characterization of urban morphology and network structure. *Remote Sens.* 11 (18), 2162.
- Small, C., Pozzi, F., Elvidge, C.D., 2005. Spatial analysis of global urban extent from dm-sp-ols night lights. *Remote Sens. Environ.* 96 (3–4), 277–291.
- Sutton, P.C., 2003. A scale-adjusted measure of “urban sprawl” using nighttime satellite imagery. *Remote Sens. Environ.* 86, 353–363.
- Tatem, A., 2017. WorldPop, open data for spatial demography. *Sci. Data.* <https://www.nature.com/articles/sdata20174.pdf>.
- Uchida, H., Nelson, A., 2009. Agglomeration Index: Towards a New Measure of Urban Concentration. World Bank, Washington, DC.
- World Bank, 2009. World Development Report, 2009: Reshaping economic geography. The World Bank, Washington, DC.
- World Bank, 2018. District Level Database for the Indonesia Urbanization Flagship Report. World Bank, Jakarta Unpublished.
- World Bank and International Monetary Fund (IMF), 2013. Global Monitoring Report 2013: Rural-Urban Dynamics and the Millennium Development Goals. The World Bank, Washington, DC.
- Zhang, Q., Seto, K.C., 2011. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sens. Environ.* 115 (9), 2320–2329.
- Zhou, N., Hubacek, K., Roberts, M., 2015. Analysis of spatial patterns of urban growth across south asia using DMSP-OLS nighttime lights data. *Appl. Geogr.* 63, 292–303.