Learners in Cities: Agglomeration and the Spatial Division of Cognition*

Marigee Bacolod
Naval Postgraduate School

Bernardo S. Blum
University of Toronto

Marcos A. Rangel
Duke University

William Strange
University of Toronto

NOV 3, 2020

Abstract

This paper makes use of a new source of psychometric data to reconsider the composition of cities, the role of sorting in urban learning, and the generation of agglomeration economies more generally. The analysis establishes that individuals in large cities tend to have greater learning capacity. The spatial distribution of learning capacity is most strongly related to the age composition of cities, specifically to the location choices of young workers with high learning capacity. This implies that observed patterns of dynamic agglomeration economies include an element of sorting of learners into cities. This, in turn, has implications for placed-based and other policies.

*We would like to acknowledge the financial support of the Desautels Centre for Integrative Thinking. Strange thanks the Social Sciences and Humanities Council of Canada. We thank *Lumosity* for providing the data used in this paper, and Nicole Ng and Robert Shafer for sharing their knowledge of the data. We are indebted to Hause Lin for the comments, suggestions, and insights he provided on earlier versions of the paper. We are also grateful for comments from seminar audiences at UQAM, UC-Irvine and Duke's Sanford School of Public Policy, and Lumosity's Science Guild series.

1 Introduction

Workers are more productive in large cities than in small cities. Whether agglomeration makes workers more productive or whether large cities are simply inhabited by different (more productive) workers is a key question. Nearly all prior research has considered one dimension of selection: sorting based on a worker's productive ability. This paper takes a very different approach to this selection issue by studying individuals' cognitive functions. In particular, it distinguishes between cognitive functions associated with accumulated knowledge, often associated with crystallized intelligence, versus cognitive functions related to the ability to solve novel problems using reasoning, e.g., fluid intelligence van Aken et al. (2016). The paper's key result is the finding that there is sorting in the second dimension, with individuals able to solve novel problems choosing disproportionately to locate in large cities. We refer to such individuals as "learners."

The analysis makes use of novel large-scale data reporting well-established psychometric measures capturing a range of cognitive functions. These measures allow us to consider the relationship between agglomeration and different aspects of cognitive ability. On the one hand, our metrics capture individuals' capacity to absorb, store, retrieve, and use codified knowledge. To illustrate, this includes the capacity to solve arithmetic problems previously learned in school. On the other hand, our metrics also capture individuals' capacity to manipulate knowledge to solve novel problems. This includes, for instance, the ability to create new conceptual frameworks. Because the data track individuals repeatedly performing the same cognitive tasks, they also allow us to observe learning directly as the ability to master these specific cognitive tasks.

The paper uses these psychometric measures to characterize individual levels of cognitive functions across space. All cognitive measures exhibit a positive relationship with city size and ZIP-code density. For larger cities and denser areas, the relationship is stronger for measures capturing the ability to manipulate knowledge to solve new problems. Furthermore, there is also a positive relationship between observed learning – the rate at which cognitive performance rises with repeated trials – and agglomeration. Moreover, although individuals' initial performances and observed learning are mildly positively correlated, the ratio of an individual's observed learning to initial performance monotonically increases with agglomeration. Taken together, these results suggest the sorting not just of the most able, but of learners, into cities.

This naturally leads to the question of how this pattern comes to be, and the answer we find is surprising. First, although an individual's education has some explanatory power, our findings hold within education groups. Second, employing Gelbach (2016)'s decomposition, we find that the observed spatial patterns in cognition are accounted for by population age differences across space. In particular, the cognitive abilities required to solve novel problems quickly decay with age, and the spatial distribution of these abilities is most strongly related to the location choices of young learners into cities.

This finding relates research on the sorting process determining the composition of cities also research on urban learning. The key urban sorting issue is the spatial pattern of worker ability. It is, perhaps, most common to equate worker education with worker productive ability (or skill). See Glaeser and Mare (2001) and Berry and Glaeser (2005), for example. Of course, education is related to but not identical to worker ability. Another common approach is to use the equality of wage to the marginal product of labor to characterize worker ability. See, for instance, Combes, Duranton, and Gobillon (2008), who estimate worker fixed effects in a wage model to examine sorting by ability. Yet another approach, less common than the previous two, is to employ the match between worker and occupation together with data on the skill requirements of jobs. In this approach, workers holding positions requiring, say, social skills are supposed to have such social skills. See Bacolod, Blum, and Strange (2009a), Bacolod, Blum, and Strange (2009b), and Bacolod, Blum, and Strange (2010), for examples of this approach.

There is more limited research on sorting based on psychometric measures. This approach builds directly on psychological research that measures key personality traits. Several such metrics have been employed, including the Armed Forces Qualification Test (an intelligence measure related to IQ), the Rotter Index (of social control), and the Rosenberg Index (of self-esteem / confidence). Among these measures, the Armed Forces Qualification Test is the most closely related to our measures of cognitive functions. However, unlike our measures, the AFQT captures a combination of cognitive abilities covering aspects of crystallized and fluid intelligence and, thus, cannot distinguish between these two aspects of cognition. Moreover, in all cases of which we are aware, the Armed Forces Qualification Test is taken from the National Longitudinal Survey of Youth (NLSY) panel. This is a small but long data set, with coverage for just a few birth cohorts and with the psychometric measures taken only once, early in an individual's life. Moreover, the relatively small number of cross sectional units imposes limitations in terms of representation of different geographic locations. As a result, these data are far from perfect for the study of the relationship between age composition and the spatial distribution of cognition. Also, the NLSY data do not allow researchers to directly observe learning as the ability to master a specific cognitive task.

There are many ways that researchers have examined the spatial dimension of learning. Much of the literature has involved specific cases. The most famous of these is Marshall (1890)'s discussion of knowledge spillovers.[1] Another approach is to consider the relationship of educational outcomes and agglomeration. Gibbons and Silva (2008), for instance, show a positive relationship. In addition, there is a vast literature on innovation and agglomeration, as discussed in Carlino and Kerr (2015). Perhaps most directly relevant to this paper is the literature on dynamic learning in wage models. Recent research of this sort suggests that larger cities make workers and firms more productive (see Baum-Snow and Pavan (2012) for evidence using U.S. data, and De La Roca and Puga (2017) for evidence using Spanish data). In particular, these studies find an important dynamic effect associated with being in a large city, i.e., the longer a worker spends in a large city the more productive she becomes. In contrast, these studies find no evidence that workers in larger cities are inherently more productive than workers in smaller cities.[2]

Our paper departs from this line of research by considering selection in learning capacity. If larger cities have a dynamic effect on productivity, i.e., they "teach" workers to be more productive over time, one should expect individuals who are better learners to sort themselves into these cities (selection in slopes). Whether better learners are also more productive (selection in levels) and the extent of the correlation between these two dimensions of individual heterogeneity are open questions.[3] Moreover, over an individual's life, high learning periods are probably those in which the individual has a relatively low stock of knowledge, i.e., at young ages. Therefore, the identification of the dynamic effect of cities

[1] See also Porter (1990) discussion of clusters.

[2] See De la Roca, Ottaviano, and Puga (2014), for evidence that confident workers, as defined by the Rosenberg index, sort disproportionately into large cities.

[3] Neither Baum-Snow and Pavan (2012) nor De La Roca and Puga (2017) allow for selection based on learning capacity.

on workers' productivity depends on whether cities are populated by the faster learners. The difficulty is that traditional data sets of workers do not measure individuals' learning abilities. Our non-traditional data confirm that learners sort themselves into cities and, thus, the estimated dynamic effects of cities must be taken with a grain of salt.

The finding that learners select into cities is also important for the light it sheds on the role of absorptive capacity in innovation. Jaffe, Trajtenberg, and Henderson (1993) show that patent citations are disproportionately local, a finding echoed in the many follow-up papers discussed in the survey by Carlino and Kerr (2015). Cohen and Levinthal (1990) note that knowledge spillovers require both generation (knowledge that could spill over) and absorption (the capabilities and incentives needed to receive the spillover). While the literature spawned by Cohen and Levinthal is concerned primarily with organizational learning, our paper's analysis of individual learning offers a new way to capture some elements of absorptive capacity.

This paper's analysis also has implications for policy. There is concern in many large cities regarding housing affordability. To the extent that this affordability issue hits young households the hardest, this paper's results suggest that affordability problems can be translated in the longer term into productivity problems. Hsieh and Moretti (2019) present a quantitative analysis of the costs of restricting the size of large cities in a somewhat related context. These costs are shown to be enormous. Of course, our results also imply that extending the urban "treatment" to workers who have not chosen it will have smaller effects than on the learners who did choose large cities.

The remainder of the paper is organized as follows. In Section 2 we start by discussing the model of cognition that is widely accepted by psychologists and neuroscientists. We aim at presenting a structure that provides context for the measures used in the data analyses that follow. Section 3 introduces the data and discusses its properties. Section 4 documents the spatial pattern of cognitive capacities. Section 5 presents a model whose predictions help us understand the sources of the cognitive sorting that we find. Section 6 concludes.

2 A Brief Discussion of Cognition

The Latin word *cognition* means "the faculty of knowing" Purves et al. (2008). Since the seminal work of Marr (1982), cognition is understood as a complex information-processing system. In particular, it is the consequence of a sequence of brain functions that i) perceive external stimuli; ii) extract key information and hold them in memory; and iii) generate thoughts and, potentially, actions. Although these multiple cognitive brain functions operate in parallel, it is pedagogically useful to describe them sequentially. Thus, one can imagine that the first step in "knowing something" consists of sensing/perceiving things external to oneself. This happens via visual, auditory, and other sensory stimulation. Then comes attention; individuals are often subjected to an enormous amount of sensory information and need to filter and focus on the relevant stimuli.[4] The next step consists of encoding the information captured and storing it to create memory. Memory is divided into short- and long-term memory. To use a computer analogy, short-term memory is like a computer's RAM, while long-term memory is analogous to a computer's hard-drive. Thus, memory is the ability to consciously remember something that has been previously captured by the individual's sensory systems. Once encoded and stored, the ability to retrieve information is the next cognitive function the brain must perform.

Supervising and regulating the brain functions just described are a set of processes that have to do with managing oneself and one's resources in order to achieve a goal. These are broadly called *executive functions* Luria (2012), Lezak et al. (2004), Miyake et al. (2000), Friedman and Miyake (2004), Miyake and Friedman (2012). In essence, the term *executive functions* is an umbrella term for the neurologically-based skills involving mental control and self regulation. Although there are different models of executive functions,[5] there is consensus that executive functions combine to produce the ability to reason, to problem solve, and to plan. One may conjecture that an individual within a heavily urbanized setting faces different demands in these dimensions than does an individual located outside of those areas.

The three basic executive functions of the brain are working memory (updating), cognitive flexibility (shifting), and inhibitory control Miyake et al. (2000). Working memory allows an individual to

[4]Incidentally, it is plausible to imagine that the urban environment around the individual affects what the individual is exposed to and relates to ones ability to focus attention.

[5]Researchers even debate the unity and the diversity of executive functions (see Friedman and Miyake (2017).

hold multiple pieces of information in mind at the same time, and to mentally work with this information. It thus allows individuals to consider alternatives, to mentally relate information to derive general principles, and to see connections between seemingly unrelated objects. Cognitive flexibility allows an individual to change the way she thinks about a situation, creating the ability to see the same scenario from different perspectives and thus to "think outside of the box." A way to think about cognitive flexibility is to understand that it is efficient, in terms of energy consumption, for the brain to operate in "autopilot" mode. In other words, once you know how to do something, thinking, i.e., questioning, and considering alternatives may get in the way of performing well. Individuals with different levels of cognitive flexibility have different costs to looking at an issue from multiple perspectives.

Finally, the third basic executive function, inhibitory control, allows an individual to control attention, thoughts, behaviors, and emotions. It is, thus, fundamental to the performance of multiple cognitive functions, including the other two basic executive functions. In order to mentally hold and manipulate thoughts one needs to inhibit other thoughts and distractions. In order to be cognitively flexible one needs to inhibit current perspectives. Executive functions, and cognitive flexibility in particular, relate to the personality traits of resilience and consciousness (see Fleming, Heintzelman, and Bartholow (2016); Genet and Siemer (2011)). Cognitive flexibility and working memory relate to creativity as well. While research has traditionally viewed intelligence and creativity as distinct and only modestly related traits (see Batey and Furnham (2006); Kaufman (2016); Runco (2007)), recent analyses show a strong correlation between measures of creativity and some executive functions, chief among them being cognitive flexibility (Nusbaum and Silvia (2011); Chen et al. (2014); Lu, Akinola, and Mason (2017); Takeuchi et al. (2011)).

The brain's executive functions produce the ability to employ information to create higher-order models of reality, a feature that has been regarded as the fundamental one in *human* cognition (Goldstein and Scheerer (1841)). Not surprisingly, an individual's executive functions closely relate to the individual's Fluid Intelligence (van Aken et al. (2016)). Recall that fluid intelligence is the ability to solve novel problems using reasoning. In contrast, crystallized intelligence is a knowledge-based ability that depends on education and acculturation (Horn and Cattell (1966)). Interestingly, differences across individuals in executive functions are almost entirely genetically driven (Friedman et al. (2008))

In sum, cognition is the output of a combination of brain functions. It seems likely that a big city individual will face a different set of cognitive challenges and opportunities than will a worker exposed to a less urban environment. In our view, these features will attract and retain workers whose cognitive capacities are most suitable to big cities. The remainder of the paper documents this phenomenon, beginning with a discussion of the data we employ.

3 Data

This paper draws on a novel large dataset on neuro-cognitive test performances of the universe of individuals in the United States who signed up for an account at the Lumosity online cognitive-training platform.[6] Individuals establish accounts at this platform (at a cost) in order to practice these cognitive functions. In the process of doing so, the platform measures and records several dimensions of an individual's cognitive performance.

The original data we have been given access to include the cognitive performances of individuals who enrolled in the platform between 2013 and 2014 and have completed at least 20 sessions of any of the four cognitive tasks further described below. For each individual we observe the individual's performance each time a session of the task is performed within this online application.[7] This score is normalized by the institution sharing the data into a rank-percentile using the universe of users and of task executions at the platform, with re-weighting designed to reproduce the 2010 U.S. Census population distribution with respect to age, education, and gender. The score transformation also minimizes trial outliers by standardizing performance using scores from beginners and more experienced users together. In all our exercises below we convert these percentile rankings into z-scores inverting a standard-Normal distribution.

Together with this information on performance, we were given access to the individual's age, gender,

[6]See `http://www.lumosity.com`

[7]Suppose the task involves performing arithmetic calculations. We refer to each calculation as a trial, and the set of trials that receives a score as a session.

education level, and the IP-address of the device used at the time of enrollment. This IP-address is mapped to a 3-digit ZIP code (ZIP3). This ZIP code is likely the area where the individual resides.[8] Finally, for each time a session is initiated the data record contextual information, including the day of the week, the time of the day, and the device being used (smartphone or computer).

3.1 Measures of agglomeration and ZIP3 characterization

Using information on the individuals' 3-digit ZIP code at the time of enrollment, we merge in data from the 2010 U.S Census to characterize the individuals' city environments. In particular, we measure the population of the metropolitan area (CBSA) in which the person lives,[9] and the population density of the 3-digit ZIP code in which the person lives (population per square mile).[10]

Appendix Figure A1 shows the distribution of individuals in our sample over our two agglomeration measures. The top panel shows a uni-modal distribution across ZIP-code population densities, with the mass of users living in areas with 500 to 3,000 inhabitants per square mile. The bottom panel shows a fairly even distribution of users across the range of sizes of metropolitan areas.[11] These figures show that our sample captures individuals from a wide variety of urban environments. Appendix Figure A2 shows that, while our two measures of agglomeration capture different aspects of the urban environment, the densest ZIP3 areas in our data tend to be in within the largest metropolitan areas. As we will discuss later in the paper, our main analytical results hold on both measures of agglomeration.

3.2 Cognitive tasks

Our data include individual performances in four cognitive tasks that map nicely into the dimensions of cognition described in the previous section. Here we briefly describe what is being captured in each task

[8]The sign-up process requires quite a few pieces of information, including credit card information, and it is natural to think that this is more likely done at home rather than at work or while in transit.

[9]We compute this by weighting population of CBSAs which overlap by the share of the zip3-area-polygon inside each CBSA.

[10]Given the growing body of evidence suggesting that agglomeration economies operate at below the metropolitan level (Rosenthal and Strange, 2003; and Ahlfeldt et al., 2015; among others), it seems sensible to carry out the analysis for a more granular geography.

[11]The exception being CBSAs with 7.5 million inhabitants.

execution.

3.2.1 Numerical Calculation

In this task, water drops are visually presented as moving down on the device screen having a (simple) arithmetic calculation within them. The individual seeks to provide the answer to the arithmetic calculation before the drop reaches the bottom of the screen. The speed at which drops show up and fall increases as the individual progresses across levels of difficulty in the task.

This task is not based on any specific psychometric test used by psychologists. Because the arithmetic expressions in the task are quite simple, it may be argued that the main cognitive function involved in this task is the ability to retrieve crystallized knowledge. This task may require some working memory, specially when the arithmetic calculation requires carryover (e.g.: 45+18). Still, in our interpretation, this task is primarily about retrieving crystallized knowledge and uses minimal, if any, ability to mentally manipulate content or cognitive flexibility.

3.2.2 Visual Span

This task consists of asking an individual to recall the location of blocks on a grid. It is based on a well-established memory test (Memory Span) that has been extensively validated in the psychology literature. In particular, this is a test of the individual's visual memory (Della Sala et al. (1997); Della Sala et al. (1999)). Visual-spatial memory allows us to remember where things are located in space, and it is necessary for navigation and mental imagery (e.g., chess positions, route to work).

Like the previous task, this one involves mostly the ability to retrieve information from memory, and once again uses minimal, if any, ability to mentally manipulate content or cognitive flexibility. Unlike the previous task, this one relies on short-term memory and, in particular, on a specific component of it: the visual-spatial part.

3.2.3 Working Memory

In this task, individuals are shown shapes in sequence, and then prompted to recall the shape shown *n* occurrences ago. It is based on a well-known test called "n-back," which has been shown to measure non-verbal working memory (Owen et al. (2005); Verhaeghen and Basak (2005)). This task primarily measures the ability to mentally hold and manipulate multiple thoughts. Such capability is what allows people to selectively and effectively control what information are retrieved from short- and long-term memory (Miyake and Friedman (2012)). The n-back task has been shown to be associated with fluid intelligence and IQ, especially when n is large (Jaeggi et al. (2010)). Because this task requires high levels of selective and sustained attention, it's been argued that it measures other aspects of executive functions as well, in particular inhibition control (Diamond (2016)). This task in fundamentally different than the previous two as it requires the mental manipulation of information.

3.2.4 Task Switching

This task presents individuals pairs of letters and numbers, and then asks the users characteristics of the letter or the number shown. By forcing the individual to switch between characteristics of the letters and numbers shown, this task is able to measure the individual's ability to switch between cognitive processes. Like the previous two tasks, this task is based on a well-known psychometric test; in this case, task switching is based on the test called "Number-letter" that has been extensively used and validated by psychologists as an instrument to measure shifting or multi-tasking and cognitive flexibility. Switching processes involve flexibly transitioning to alternative task-set representations in the prefrontal cortex (Miyake and Friedman (2012)). It helps people shift attention, reconfigure perceptual and responses as the context changes, and is a critical process that allows people to meet changing/shifting demands in their everyday lives (Monsell (2003); Rogers and Monsell (1995)). Like "Working memory", this task requires the mental manipulation of information as well.

3.3 Analytical sample and descriptive statistics

Given the research question, our analysis sample is limited to users of this platform with a relatively extensive history of practice. That is, individuals who practiced a given cognitive task at least 20 times after enrolling in the platform. We further restrict our analysis to those who finished their 20 sessions within 2 years after the initial sign-up date. In order to make our results more relevant to the existing economics literature, which focuses on individuals active in the labor force, we also restrict our working sample to individuals between the ages of 22 and 59. Panel A of Table 1 reports summary statistics on this sample. From the sample sizes it is easy to observe that there are differences in popularity across the tasks, particularly with working memory tasks engaging less individuals. While there are differences in the characteristics of individuals engaging in these different tasks, all results we present below are unaffected by restricting measures to a subset of individuals engaging in all of them.

When considering the overall U.S. population, one characteristic of our sample that immediately stands out is that individuals are significantly more educated than the average working-age American, with 64% to 67% having a college degree or more. That more educated individuals are more likely to enroll in cognitive training exercises is perhaps not surprising. Given our research question, the dimension over which sample representativeness particularly matters is along measures of agglomeration. In Panel A of Appendix Figure A3 we portray education (% college or more) characteristics of our sample (dashed lines) compared with the Census (solid lines) across different values of ZIP density. Once again its is easy to see that our sample drew disproportionately fewer less educated individuals than their proportions in the population. The same figure also shows that the relationship between the share of college educated individuals and ZIP density is similar in our sample and in the Census. That is, the slopes of the two lines in Panel A very similar. Thus, even though our sample is disproportionately composed of college graduates, this is equally the case in small and large cities relative to what is found in the census. We find similar results when we use CBSA population as the agglomeration measure.

Panels B carries out a similar analysis for the gender composition of our sample. In particular, it ploys gender (% male) characteristics of our sample (dashed lines) compared with the Census (solid lines) across ZIP density. Our sample drew disproportionately fewer males, in particular in less dense are. As we will discuss later, gender is not an important driver of our results and will not be the focus of our analysis.

Finally, Panel C of the same figure indicates that younger individuals in our sample are slightly more represented in denser areas than in the population, even though the overall pattern of the younger inhabiting denser areas is present both in the Census and in our sample. We will have more to say about this dimension of selection when we interpret and discuss our findings.

Our analysis will focus on a sub-sample of individuals with at least a college degree. This is because there is a well documented relationship between cognitive functions and education attainment. The focus on college graduates allows us to obtain more precise estimates because of the sample size, as well as allow a cleaner interpretation of the estimates.[12] Panel B of Table 1 reports descriptive statistics on our final analytical sample. The average user is in their early 40's, with only between 16% to 24% of individuals under 30, and females continue to be over-represented in our sample, with 43% to 45% being composed of males.

4 Cognition across space

4.1 Average performance

In this section, we characterize the spatial patterns of the four cognitive functions measured in our data. Figure 1 depicts the nonparametric cognition-agglomeration relationship. More precisely, this figure shows the average performance of individuals against our two measures of agglomeration.[13] The figure shows a positive relationship between all four metrics of cognitive function and both metropolitan population and zip-code population density. Interestingly, in both cases there are clearly steeper relations between cognition and agglomeration as we consider measures more closely related to the brain's executive functions. These steeper relations are highlighted by the bifurcations apparent in both figures. When agglomeration is measured by metropolitan population, the bifurcation happens around the population

[12]The key qualitative results persist for the entire sample.

[13]We trim the top and bottom 1% of the performance distribution to avoid ceiling and flooring effects.

level of Memphis, TN; it happens around the zip-code density of Stamford-CT (zip3 069). Above these two points, average performance rises more rapidly for the measures of higher brain functions, Working Memory and Task Switching, showing that in large cities we tend to see individuals with relatively higher levels of the brain functions associated with solving new problems through reasoning.

To confirm that the positive slopes depicted in the previous figure are statistically significant, we estimate parsimonious linear relationships between cognitive performances and agglomeration. These models take the form:

$$\bar{S}_i^c = \alpha_0^c + \alpha_1^c AGG_i + \epsilon_i^c \tag{1}$$

where (c, i) index cognitive task and individual respectively, $\bar{S}_i^c$ captures average individual *i*'s performance across 20 sessions of cognitive task c, and AGG_i represents one of our agglomeration measures.

Panel A of Table 2 shows results and confirms that the estimated relationships are statistically significant and that the magnitudes of the effects are large.[14] In particular, doubling the ZIP density is associated with an increase in the average performance on the Working Memory, Brain Shift, or Spatial Memory tasks equivalent to 10% of the raw college-to-high-school gap of performance in the same tasks. The slopes for Numerical Calculations are notably smaller, and not statistically significant when agglomeration is measured by CBSA population.[15]

4.2 Comparative advantage in executive function

The analysis thus far suggests that, in addition to having higher absolute levels of all cognitive functions, individuals in denser and larger areas have a comparative advantage on the executive cognitive functions of the brain. To consider comparative advantage, however, we must deal with the fact that users of the platform choose which tasks to practice. Thus, to study the comparative advantage of individuals across tasks, we focus on a sub-sample where individuals engaged in all four tasks. Table A2 in the

[14]For completeness, in the Appendix (Table A1) we show that most of these findings hold on the sub-sample of individuals with less than college as well, even though the results are understandably noisier.

[15]As seen in the previous figure and table, throughout the analysis we find similar patterns when we use Zip density and CBSA population. To conserve space, henceforth we report the results using CBSA population as the measure of agglomeration in the appendix only.

Appendix shows summary statistics on this sub-sample of 27,109 individuals. Appendix Figure A5 depicts, non-parametrically, the cognition-agglomeration relationship in this sub-sample and confirms the same patterns as in the figure using our analytical sample.

We study how an individual's comparative advantage in a task varies with agglomeration using a conditional regression analysis. In particular, we measure how performance in three tasks relate to agglomeration after netting out the individual's performance in the Numerical Calculation task. Figure 2 shows the average conditional performance across sessions against the ZIP-density measure of agglomeration. It confirms that, indeed, individuals tend to have better performances in the higher cognitive functions relative to their performance in the Numerical Calculation task.

Again, we estimate linear relationships between cognitive performance and agglomeration, this time controlling for the individuals' performances in the Numerical Calculation task, and check the statistical significance of the correlation between cognition and agglomeration. Specifically, we estimate the following equation:

$$\bar{S}_i^c = \gamma_0^c + \gamma_1^c \bar{S}_i^{NC} + \gamma_2^c AGG_i + \upsilon_i^c \tag{2}$$

Like in the previous equation, c indexes cognitive task, but now c ={Visual Span, Working Memory, Task Switching}. $\bar{S}_i^{NC}$ represents individual *i*'s average performance in 20 sessions of the Numerical Calculation task. Panel B of Table 2 confirms with regressions the positive and significant slopes of the curves in the figure.

4.3 The evolution of cognitive performance: observed learning

Because we observe each individual in our sample performing the same cognitive task 20 times, we can study how individual performances change as the individual performs the task more times. For this, we revert to our analytical sample.

Perhaps not surprisingly, Figure 3 shows that performances in all tasks improve over sessions. This improvement, a kind of observed learning, happens at diminishing rates and, by the twentieth session performances have pretty much stabilized. Moreover, the different cognitive tasks have different scopes for improvement over sessions, with the two tasks measuring the higher order cognitive functions having higher growth over sessions, on average.

How should these results be interpreted? One interpretation is that, by performing the tasks on the online platform, individuals are actually improving their cognitive functions. In the case of the two executive functions, individuals would be learning how to be better learners. It is also possible that the performance improvement over sessions reflects individuals learning to handle the mechanics of how to play the task, i.e., learning-by-doing. Of course, these possibilities are not mutually exclusive. In any event, the patterns in Figure 3 reveal learning. Given the previous evidence showing that individuals in larger and denser areas have more of the cognitive functions that allow solving new problems via reasoning, it is natural to ask whether these individuals are actually better at figuring out the cognitive tasks over sessions.

To study this, we compare the relationship between agglomeration and individuals' average cognitive performance over their first and last five sessions. Figure 4 shows these results. Two features in this figure are striking. First, there is a monotonic and strong relationship between agglomeration and cognitive performance in the individuals' initial performances. That is, individuals in denser cities start with better performances. Second, and perhaps more interesting, the performance-agglomeration relationship becomes steeper for the individuals' later performances. Thus, not only do individuals in denser areas start with higher performance, but they also improve their performances at higher rates across sessions.

Confirming these findings, Figure 5 shows the non-parametric relationship between performance in the individual's last five sessions (sessions 16 to 20) and agglomeration, after controlling for the individual's average performance in her first five sessions. More precisely, we regress average performance in the final 5 sessions on average performance in first 5 sessions and, then, we plot the residual of this regression against agglomeration. Quite clearly, individuals in denser areas tend to have larger performance improvements over sessions, even controlling for the fact that they start from higher levels.

In summary, our analysis thus far shows that individuals in denser and larger areas have better cognitive functions. Moreover, these individuals are learners in the sense that they display higher levels of the brain functions associated with the ability to figure out new things through reasoning. Indeed, we confirm that these individuals are better at figuring out how to improve their performances in the cognitive tasks.

5 A model of learners in cities

How does the pattern come to be where large cities attract more individuals with the ability to figure out new things? To answer this question, we build a model of city choice with sorting based on learning capacity.

The model has two key assumptions, both with ample empirical support. First, learning capacity decays with an individual's age. Second, individuals carry their human capital with them even if they move urban settings. The first assumption has strong support in the psychology and neuroscience literatures. Indeed, in addition to education, the other characteristic that strongly affects individual's cognitive functions is age. In particular, cognitive functions decay steadily with age, starting as young as the early 30's. Moreover, the evidence is that higher-order functions, the ones we find disproportionately more in cities, are the ones with sharper and earlier declines (see Hartshorne and Germine (2015)). Figure 6 confirms, in our sample, that task performances decay quite strongly with age. Figure 7 shows that learning over sessions also decays with age in our sample. The assumption that individuals carry their human capital with them even if they move urban settings is also well-supported (see, for instance, Glaeser and Mare (2001) or De La Roca and Puga (2017)).

5.1 Basics

The literature on urban sorting has largely focused on models where sorting is based on the worker's productive ability. Glaeser and Mare (2001) and Combes, Duranton, and Gobillon (2008) are seminal papers in this literature. Wang (2016) is a more recent contribution. In Baum-Snow and Pavan (2012) and De La Roca and Puga (2017), wage growth is interpreted as improvements in ability due to learning, which cities foster.[16]

Our study works with a model of city choice that instead focuses on sorting based on learning capacity. Workers are heterogeneous in several ways. Some are learners, while some are not. Some are young, while some are old. Some of the old learners will have learned while they are young. Others will not have learned. Finally, workers have idiosyncratic tastes for the amenities offered by different cities.

There are two cities that a worker might choose, a big city and a small city. Locating in a city imposes costs c^B and c^S, with higher costs in the big city: $c^B > c^S$. A young learner earns wage w^{LB} or w^{LS}, depending on the city size chosen. A young nonlearner earns wage w^{NB} or w^{NS}. An old nonlearner continues to earn w^{NB} or w^{NS}. An old learner who lived in a small city while young is assumed not to have learned during his/her youth, and so such a worker continues to earn w^{LB} or w^{LS}. And old learner who lived in a big city while young is assumed to have learned, and this worker now enjoys an additional ω in wage. This is in addition to the wages described above. Finally, each worker has an idiosyncratic taste for the two differently sized cities. Let ϕ_i denote the additional utility in the larger city for worker i, with $f(-)$ and $F(-)$ denoting ϕ_i's probability density and cumulative distribution functions.

The location choice of a worker of a given type will depend on lifetime utility. For simplicity, we suppose that workers live one period as young and one additional period as old, with no discounting. For each type of worker (e.g., young learner, young nonlearner, etc...), there will be a critical big city amenity value, ϕ^{*z}, such that a worker of type-z with big city taste ϕ^{*z} will be exactly indifferent between big and small cities. Workers with stronger tastes, $\phi_i > \phi^{*z}$ will strictly prefer the big city, while workers with weaker tastes, $\phi_i < \phi^{*z}$ will strictly prefer the small city.

The solution for the composition of cities of the two sizes will be determined by critical amenity level ϕ^{*z} for each type. Beginning with the old individuals who are not learners, the equal utility condition defining the critical amenity level is:

$$(w^{NB} - c^B) + \phi^{*ON} = (w^{NS} - c^S) \tag{3}$$

[16] De la Roca, Ottaviano, and Puga (2014) is an exception, considering sorting on self-confidence, as measured by the Rosenberg Index.

This implies:

$$\phi^{*ON} = (w^{NS} – c^{S}) - (w^{NB} - c^{B}) \tag{4}$$

Consider now an old learner who inhabited a small city as a young worker and so did not learn. For this worker, solving the equal utility condition for ϕ^{*OLS} gives

$$\phi^{*OLS} = (w^{NS} – c^{S}) - (w^{NB} - c^{B}) = \phi^{*ON} \tag{5}$$

Similarly, an old worker who did learn earns the extra premium of ω in whatever city is chosen in the second period. This gives:

$$\phi^{*OLB} = (w^{NS} – c^{S}) - (w^{NB} - c^{B}) = \phi^{*ON} = \phi^{*OLS}. \tag{6}$$

Turning to the young workers, a young nonlearner faces exactly the same situation as an old non-learner. The critical amenity level is again:

$$\phi^{*YN} = (w^{NS} – c^{S}) - (w^{NB} - c^{B}) = \phi^{*ON} = \phi^{*OLS} = \phi^{*OLB}. \tag{7}$$

For a young learner, however, the situation is different. Such a worker chooses between a big city utility of ($w^{NB} - c^{B} + V^{OB}$) and a small city utility of (w^{NS}–$c^{S} + V^{OS}$),where V^{OB} and V^{OS} denote the utility of an old worker depending on the city chosen while young. This gives:

$$\phi^{*YN} = (w^{NS} – c^{S}) - (w^{NB} - c^{B}) + (V^{O^{S}} - V^{O^{B}} = (w^{NS} – c^{S}) - ((w^{NB} - c^{B}) = \omega, \tag{8}$$

since the difference between utility as an old worker, (V^{OS} - V^{OB}), equals the wage premium, ω. This implies $\phi^{*YL} < \phi^{*YN}$= ϕ^{*ON} = $\phi^{*OL/1S}$= $\phi^{*OL/1B}$. The two city sizes then have compositions determined by $F(-)$.

This characterizes the equilibrium choices of workers of various types. Young learners require a smaller level of amenities to choose the big city since they, unlike all of the other worker types, have the additional learning incentive to locate in the large city. The key result is: A large city will have a greater share of young learners than will a small city (Proposition 1).

5.2 Extensions

In the model sketched above, young learners choose the large city because they are more capable of learning, and so they obtain more valuable learning from the big city. Now suppose that young workers also choose their education levels. Suppose that only learners can acquire education and that educated workers incur costs, which gives them an additional wage premium. If the education costs are distributed across workers according to some distribution, then there will be a critical level below which a given learner will choose to acquire education. Suppose that the education costs are distributed independently of the idiosyncratic tastes discussed above. An immediate corollary of Proposition 1 is that the big city will contain more educated workers. This is a consequence of the attraction of learners to the big city. It holds even without assuming a complementarity between worker education and city size in the production process.

One way to obtain such a complementarity would be to assume that ability has more value in the large city. Suppose, for instance, that all wages are multiplied by $\alpha > 1$ in large cities, including the additional wage associated with learning. This would introduce the more traditional form of sorting, with high ability workers choosing to move to and remain in large cities because of the high wages that their productivity commands. This would lead the learned (old learners who occupied big cities in youth) to remain in large cities. Similarly, if there were relocation costs, then some older individuals would presumably remain in large cities even without a learning benefit and some young workers would choose not to learn, since they would be setting themselves up for a costly relocation to a small city when they became old.

5.3 Evidence on the model's predictions

The main message of the model developed in the last sections is that living in a city is an investment and, like any investment, it depends on the expected present value of its return. Specifically, the return to living in a city depends on the individual's cognitive functions. These vary across individuals but, in general, increase with education and decrease with age. The return to living in a city will depend directly on the individual's age as well, through the present value calculation. Thus, within education groups, age should be the main mediator of the relationship between agglomeration and cognitive functions.

To take this idea to the data, we start by studying the effect of controlling for individuals' age on the relationship between agglomeration and cognitive functions. More precisely, we estimate:

$$\bar{S}_i^c = \bar{X}_i'\beta^c + \theta^c AGG_i + \epsilon_i^c \tag{9}$$

where, as before, (c, i) index cognitive task and individual, respectively, and $\bar{S}_i^c$ captures individual i's performance across 20 sessions of cognitive task c. The vector $\bar{X}$ contains the individuals' age, gender, and education variation within college graduates. It also contains contextual information averaged across the twenty sessions, such as the day of the week and time of the day the session was performed, and the device used (smartphone or computer). This contextual information may proxy for unobserved individual characteristics such as income and occupation (performing the task during business hours or during weekends).

Table 2 shows results for each cognitive task (reported in separate columns). In Panel C, we report the coefficient for the same relationship with agglomeration, now conditional on the individual's age, gender, education variation within college graduates, and average circumstance across 20 sessions. In most cases, the controls make the relationship between cognitive performance and agglomeration turn from positive to zero.

At the bottom of Panel C we also report a decomposition that attributes to each covariate in $\bar{X}_i$ the percentage of the observed change in the agglomeration coefficient that is due to that element of the control vector (see Gelbach (2016)). In all cases, individuals' age is what mostly account for the positive relationship between cognitive performance and agglomeration. This is consistent with the model's prediction that age intermediates the connection between cognitive function and agglomeration. In other words, on average, denser areas have individuals with better cognitive performances because they have disproportionately more young residents.

However, Proposition 1 indicates that the sorting of high cognitive performance individuals into cities should be heterogeneous across age groups. In particular, the sorting should be stronger for younger individuals and decline with individuals' age. To check for this, Table 3 shows the difference in the relation between cognition and agglomeration for individuals younger than 45 years old versus those older than 45 by interacting an indicator function for the younger group with our agglomeration measure. In all cases, sorting is significantly different for older and younger individuals, with relatively stronger positive sorting dominant among the younger individuals. Figure 8 shows, semi-parametrically, the agglomeration effects by age. This figure confirms that high cognitive performance young individuals sort themselves into cities, but this sorting decays and even reverses itself with individuals' age.

Finally, Tables 4 and 5 show similar results when we look at the sorting of better learners into denser areas by examining the mastering of these tasks. Particularly relevant are the agglomeration sorting effects along high-order brain functions (task-switching capabilities).

6 Conclusion

Cities are different than other places, and not just because they have more people. They also have a more productive workforce. This has been documented in various ways. Urban workers have higher wages, implying higher marginal productivities. They have higher levels of various cognitive and social skills, as can be seen from the skill requirements of urban occupations. They have high levels of general intelligence, as seen in standard psychometric tests. They are also more educated, which is correlated with productivity.

This paper has taken a different approach to understanding the spatial division of ability. It does by analyzing new data on work-force abilities that characterizes different sorts of cognitive performance, specifically executive function, crystallized intelligence, and fluid intelligence. Larger cities have higher levels of both sorts of intelligence, but the relationship is stronger for fluid intelligence. And the relationship is nonlinear, with the relationship between agglomeration and fluid intelligence becoming strongest for the largest cities and most dense locations.

These relationships are present both for psychometric measures and for the actual learning that can be observed from an individual's repetition of the tasks generating the psychometric scores. Results are consistent across these two approaches, with cities containing disproportionate shares of workers showing high learning capacity. The sorting of learners into large cities is shown to arise primarily from the sorting of young workers into cities. Education (particularly postgraduate) plays a role as well, but a relatively small one.

In addition to illuminating the composition of cities, the paper's results also shed light on urban learning, and thus on the generation of agglomeration economies. It is not just the most productive workers who sort into cities; workers with high capacities to learn do so as well. This means that observed learning is a product of both the urban environment (as is conventional in considering knowledge spillovers) and also on the high learning capacities of its residents (a new finding).

This conclusion bears on various sorts of place-based policies. Land use regulation can impede access to agglomeration, and the lost benefits of agglomeration can be a huge welfare cost, as demonstrated by Hsieh and Moretti (2019). Conversely, the increased benefits of agglomeration economies can be a welfare gain associated with place-based policies, as is seen in the survey by Neumark and Simpson (2015). In both cases, a causal measure of the effects of agglomeration is required. Previous discussions of sorting, which obviously compromises causality, consider sorting by ability or productivity. Our analysis, in contrast, considers sorting by learning capacity, the presence of which would also be problematic for policy analysis.

References

Bacolod, Marigee, Bernardo S Blum, and William C Strange. 2009a. "Skills in the city." *Journal of Urban Economics* 65 (2):136–153.

———. 2009b. "Urban interactions: Soft skills versus specialization." *Journal of Economic Geography* 9 (2):227–262.

———. 2010. "Elements of skill: traits, intelligences, education, and agglomeration." *Journal of Regional Science* 50 (1):245–280.

Batey, Mark and Adrian Furnham. 2006. "Creativity, intelligence, and personality: A critical review of the scattered literature." *Genetic, social, and general psychology monographs* 132 (4):355–429.

Baum-Snow, Nathaniel and Ronni Pavan. 2012. "Understanding the city size wage gap." *The Review of economic studies* 79 (1):88–127.

Berry, Christopher R and Edward L Glaeser. 2005. "The divergence of human capital levels across cities." *Papers in regional science* 84 (3):407–444.

Carlino, Gerald and William R Kerr. 2015. "Agglomeration and innovation." In *Handbook of regional and urban economics*, vol. 5. Elsevier, 349–404.

Chen, Qunlin, Wenjing Yang, Wenfu Li, Dongtao Wei, Haijiang Li, Qiao Lei, Qinglin Zhang, and Jiang Qiu. 2014. "Association of creative achievement with cognitive flexibility by a combined voxel-based morphometry and resting-state functional connectivity study." *Neuroimage* 102:474–483.

Cohen, Wesley M and Daniel A Levinthal. 1990. "Absorptive capacity: A new perspective on learning and innovation." *Administrative science quarterly* :128–152.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. "Spatial wage disparities: Sorting matters!" *Journal of urban economics* 63 (2):723–742.

De la Roca, Jorge, Gianmarco IP Ottaviano, and Diego Puga. 2014. "City of dreams." .

De La Roca, Jorge and Diego Puga. 2017. "Learning by working in big cities." *The Review of Economic Studies* 84 (1):106–142.

Della Sala, S, C Gray, A Baddeley, and L Wilson. 1997. "Visual Patterns Test: A test of short-term visual recall." *Thames Valley Test Company* 40.

Della Sala, Sergio, Colin Gray, Alan Baddeley, Nadia Allamano, and Lindsey Wilson. 1999. "Pattern span: a tool for unwelding visuo–spatial memory." *Neuropsychologia* 37 (10):1189–1199.

Diamond, Rebecca. 2016. "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000." *American Economic Review* 106 (3):479–524.

Fleming, Kimberly A, Samantha J Heintzelman, and Bruce D Bartholow. 2016. "Specifying associations between conscientiousness and executive functioning: Mental set shifting, not prepotent response inhibition or working memory updating." *Journal of personality* 84 (3):348–360.

Friedman, Naomi. P. and Akira Miyake. 2004. "The Relations Among Inhibition and Interference Control Functions: A Latent-Variable Analysis." *Journal of experimental psychology: General* 133 (1):101–135.

Friedman, Naomi P and Akira Miyake. 2017. "Unity and diversity of executive functions: Individual differences as a window on cognitive structure." *Cortex* 86:186–204.

Friedman, Naomi. P., Akira Miyake, Susan E. Young, John. C. DeFries, Robin P. Corley, and John K. Hewitt. 2008. "Individual differences in executive functions are almost entirely genetic." *Journal of experimental psychology: General* 137 (2):201.

Gelbach, Jonah B. 2016. "When do covariates matter? And which ones, and how much?" *Journal of Labor Economics* 34 (2):509–543.

Genet, Jessica J and Matthias Siemer. 2011. "Flexible control in processing affective and non-affective material predicts individual differences in trait resilience." *Cognition and emotion* 25 (2):380–388.

Gibbons, Stephen and Olmo Silva. 2008. “Urban density and pupil attainment.” *Journal of Urban Economics* 63 (2):631–650.

Glaeser, Edward L and David C Mare. 2001. “Cities and skills.” *Journal of labor economics* 19 (2):316–342.

Goldstein, Kurt and Martin Scheerer. 1841. *Abstract and concrete behavior an experimental study with special tests*. Psychological monographs 53, no. 2.

Hartshorne, Joshua K and Laura T Germine. 2015. “When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span.” *Psychological science* 26 (4):433–443.

Horn, John L and Raymond B Cattell. 1966. “Refinement and test of the theory of fluid and crystallized general intelligences.” *Journal of educational psychology* 57 (5):253.

Hsieh, Chang-Tai and Enrico Moretti. 2019. “Housing constraints and spatial misallocation.” *American Economic Journal: Macroeconomics* 11 (2):1–39.

Jaeggi, Susanne M, Barbara Studer-Luethi, Martin Buschkuehl, Yi-Fen Su, John Jonides, and Walter J Perrig. 2010. “The relationship between n-back performance and matrix reasoning—implications for training and transfer.” *Intelligence* 38 (6):625–635.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson. 1993. “Geographic localization of knowledge spillovers as evidenced by patent citations.” *the Quarterly journal of Economics* 108 (3):577–598.

Kaufman, James C. 2016. *Creativity 101*. Springer publishing company.

Lezak, Muriel Deutsch, Diane B Howieson, David W Loring, Jill S Fischer et al. 2004. *Neuropsychological assessment*. Oxford University Press, USA.

Lu, Jackson G, Modupe Akinola, and Malia F Mason. 2017. ““Switching On” creativity: Task switching can increase creativity by reducing cognitive fixation.” *Organizational Behavior and Human Decision Processes* 139:63–75.

Luria, Aleksandr Romanovich. 2012. *Higher cortical functions in man*. Springer Science & Business Media.

Marr, D. 1982. "Vision: A Computational Approach (San Fr."

Marshall, Alfred. 1890. *Principles of Economics, by Alfred Marshall*. Macmillan and Company.

Miyake, Akira and Naomi P Friedman. 2012. "The nature and organization of individual differences in executive functions: Four general conclusions." *Current directions in psychological science* 21 (1):8–14.

Miyake, Akira, Naomi P Friedman, Michael J Emerson, Alexander H Witzki, Amy Howerter, and Tor D Wager. 2000. "The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis." *Cognitive psychology* 41 (1):49–100.

Monsell, Stephen. 2003. "Task switching." *Trends in cognitive sciences* 7 (3):134–140.

Neumark, David and Helen Simpson. 2015. "Place-based policies." In *Handbook of regional and urban economics*, vol. 5. Elsevier, 1197–1287.

Nusbaum, Emily C and Paul J Silvia. 2011. "Are intelligence and creativity really so different?: Fluid intelligence, executive processes, and strategy use in divergent thinking." *Intelligence* 39 (1):36–45.

Owen, Adrian M, Kathryn M McMillan, Angela R Laird, and Ed Bullmore. 2005. "N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies." *Human brain mapping* 25 (1):46–59.

Porter, Michael E. 1990. "The competitive advantage of nations." *Harvard business review* 68 (2):73–93.

Purves, D, EM Brannon, R Cabeza, SA Huettel, KS LaBar, ML Platt, and MG Woldorff. 2008. "Principles of cognitive neuroscience (Vol. 83). Sunderland, MA."

Rogers, Robert D and Stephen Monsell. 1995. "Costs of a predictible switch between simple cognitive tasks." *Journal of experimental psychology: General* 124 (2):207.

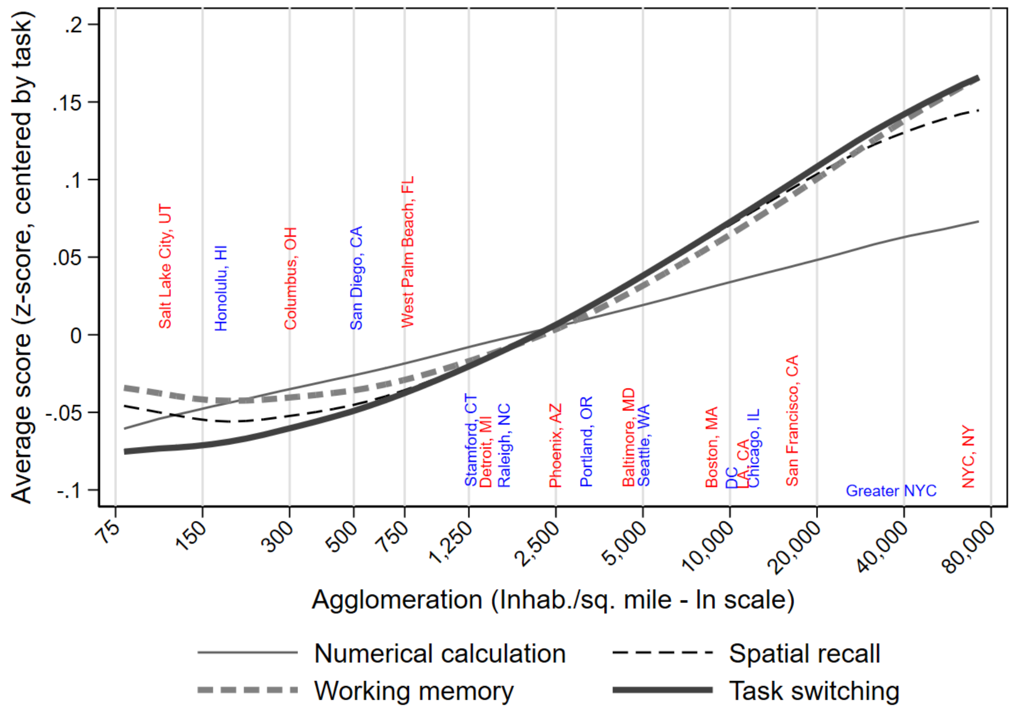
Runco, Mark A. 2007. “A hierarchical framework for the study of creativity.” *New Horizons in Education* 55 (3):1–9.

Takeuchi, Hikaru, Yasuyuki Taki, Hiroshi Hashizume, Yuko Sassa, Tomomi Nagase, Rui Nouchi, and Ryuta Kawashima. 2011. “Failing to deactivate: the association between brain activity during a working memory task and creativity.” *Neuroimage* 55 (2):681–687.

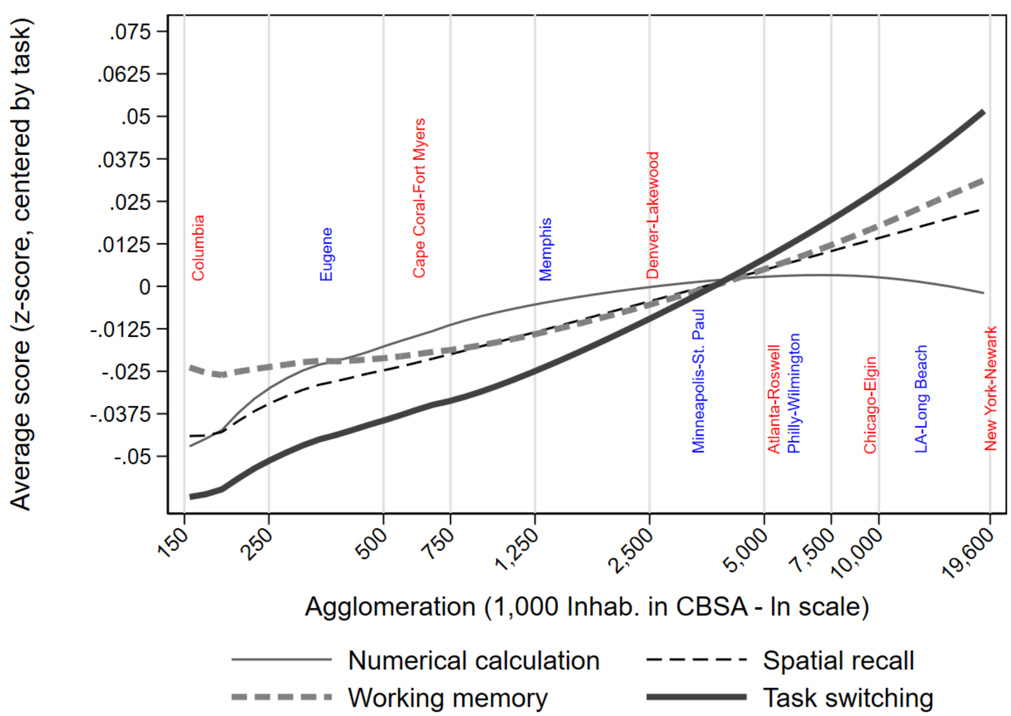
van Aken, Loes, Roy PC Kessels, Ellen Wingbermühle, William M. van der Veld, and Jos IM Egger. 2016. “Fluid intelligence and executive functioning more alike than different?” *Acta Neuropsychiatr* 28:31 – 37.

Verhaeghen, Paul and Chandramallika Basak. 2005. “Ageing and switching of the focus of attention in working memory: Results from a modified N-Back task.” *The Quarterly Journal of Experimental Psychology Section A* 58 (1):134–154.

Wang, Zhi. 2016. “Wage growth, ability sorting, and location choice at labor-force entry: New evidence from US Census data.” *Journal of Urban Economics* 96:112–120.



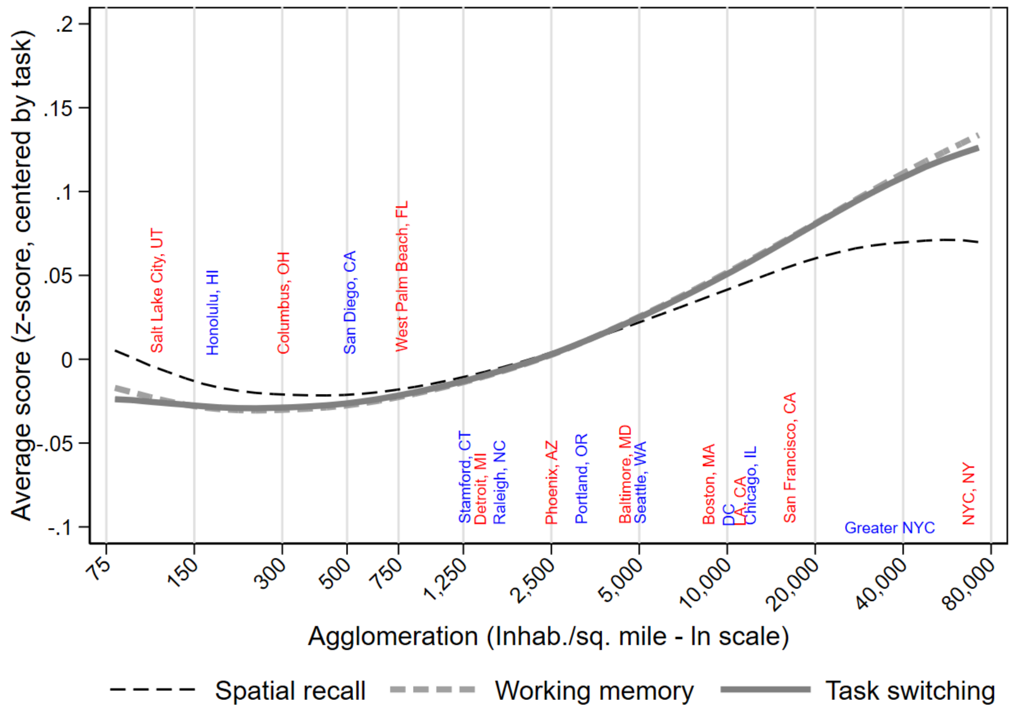
A.



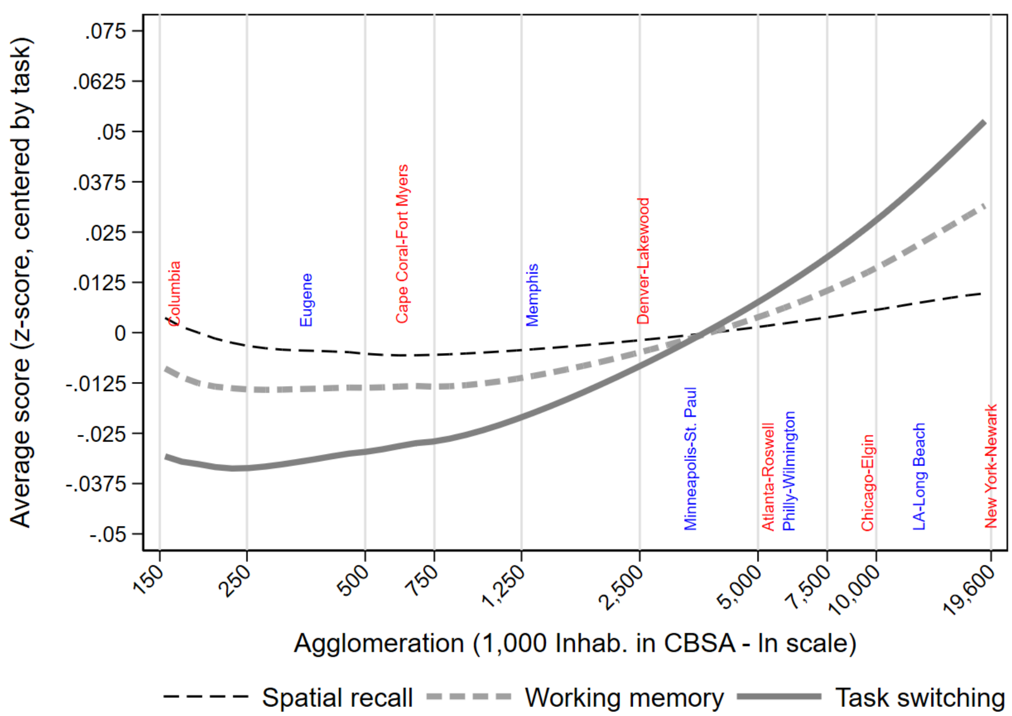
B.

Figure 1: Average performance in tasks and agglomeration, College+ sample

Notes: Sample includes users with college or more as described in text. Agglomeration measures refer to individuals' CBSA population and ZIP3-level density.



A.



B.

Figure 2: Relative performance in tasks (relative to Numerical Calculation) and agglomeration, College+ sample

Notes: Sample includes users with college or more as described in text. Agglomeration measures refer to individuals' CBSA population and ZIP3-level density.

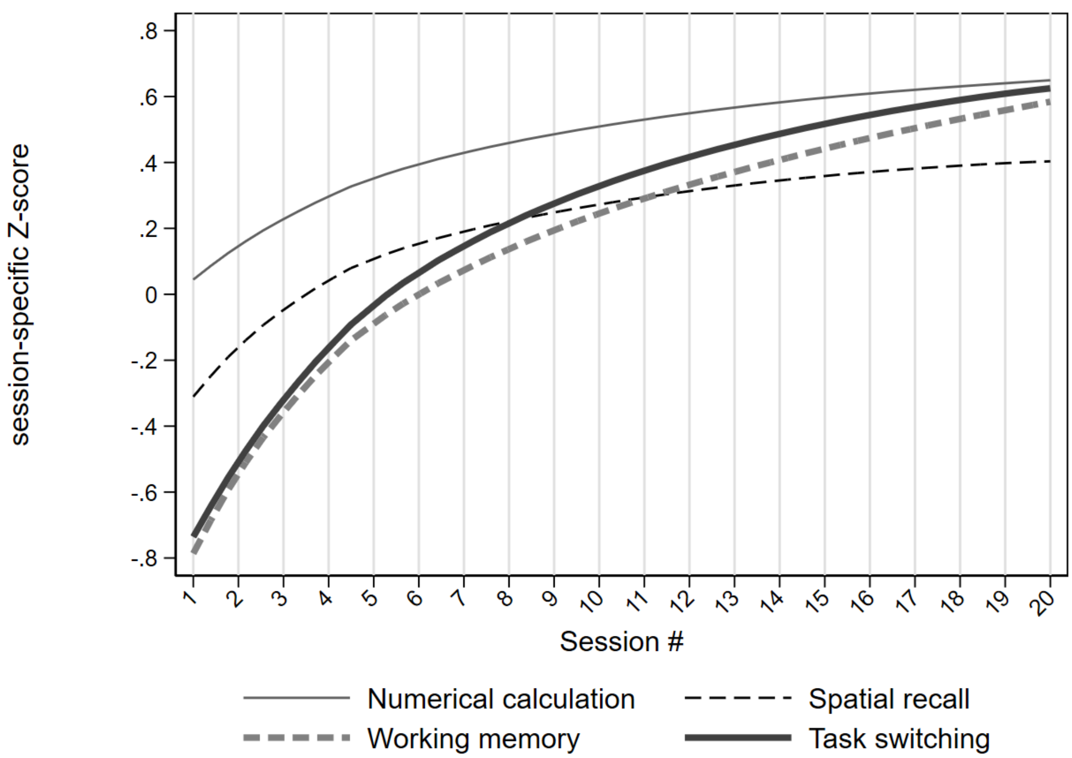


Figure 3: Evolution of performance across sessions, College+ sample

Notes: Sample includes users with college or more as described in text.

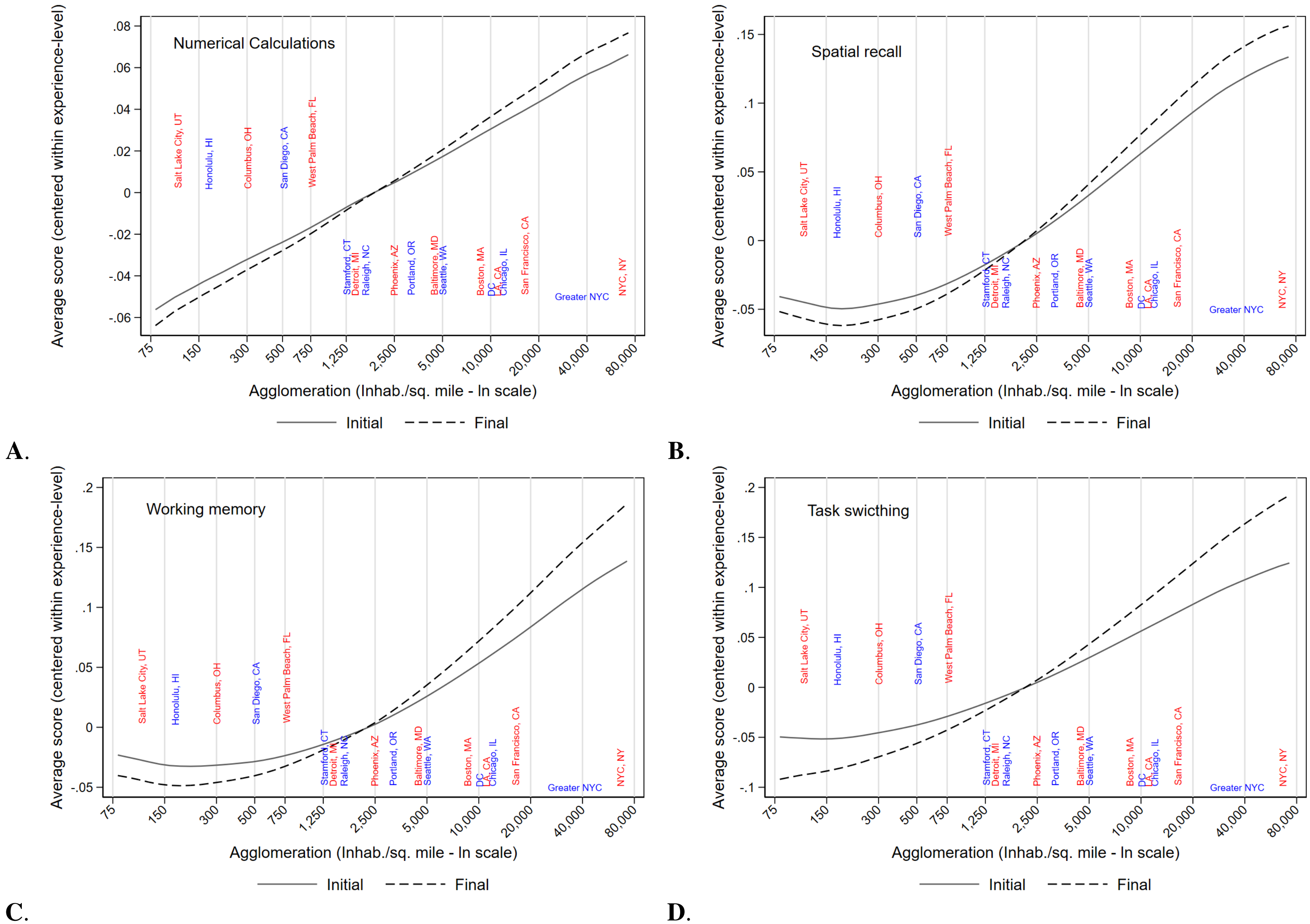


Figure 4: Initial and final performance and agglomeration by task, College+ sample

Notes: Sample includes users with college or more as described in text. Agglomeration measures refer to individuals' ZIP3-level density.

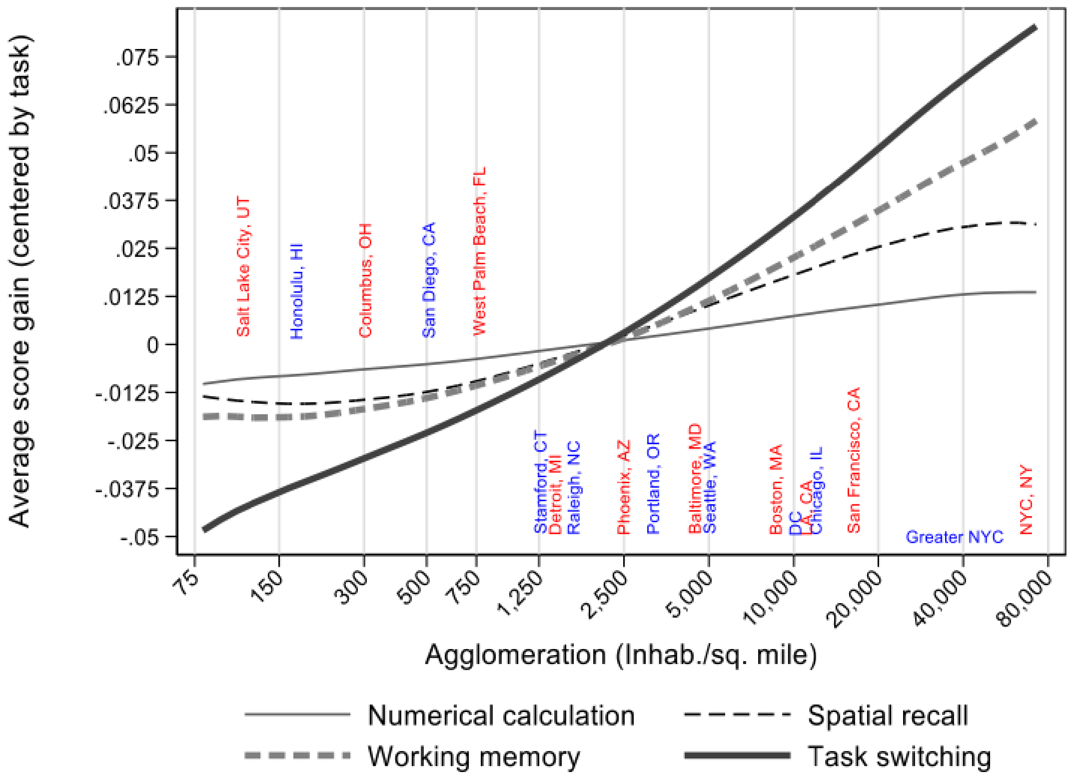


Figure 5: Gains in performance and agglomeration, College+ sample

Notes: Sample includes users with college or more as described in text. Agglomeration measures refer to individuals' ZIP3-level density. Figure plots residuals from regression of average performance in the final 5 sessions on average performance in first 5 sessions.

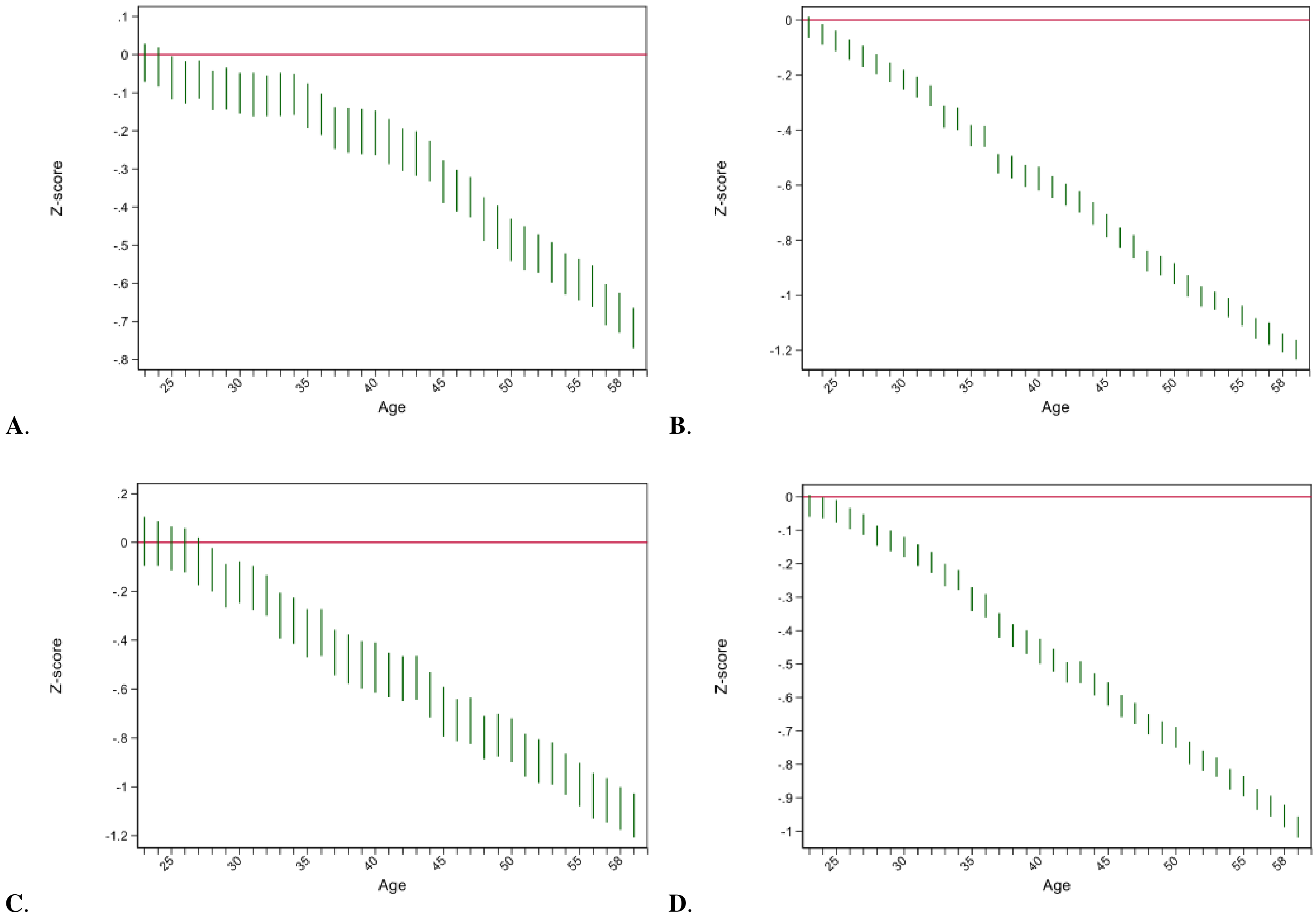


Figure 6: Average performance and age, College+ sample

Notes: Numerical calculation (Panel A), Spatial recall (Panel B), Working memory (Panel C), Task switching (Panel D). Sample includes users with college or more as described in text.

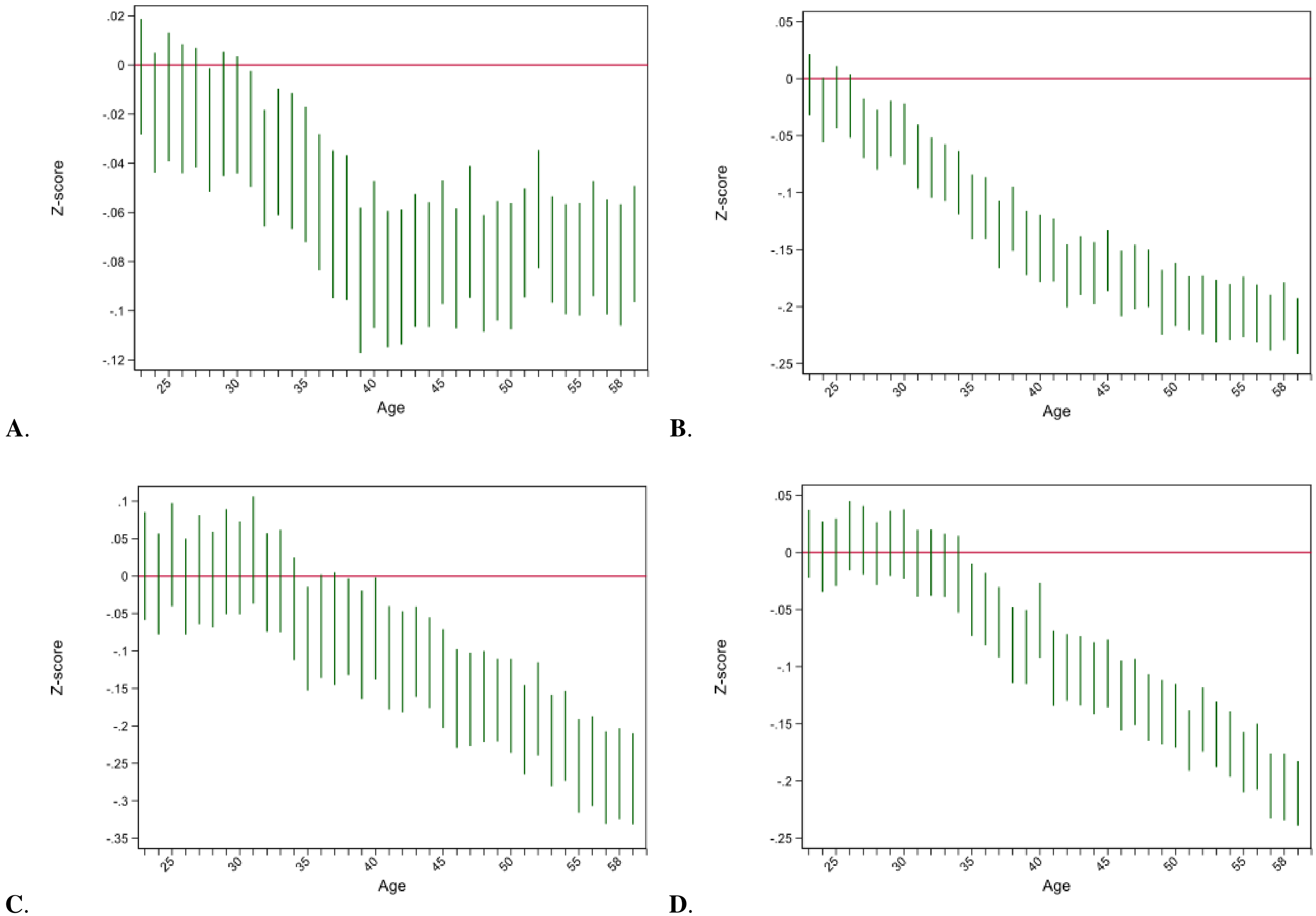


Figure 7: Gains in performance and age, College+ sample

Notes: Numerical calculation (Panel A), Spatial recall (Panel B), Working memory (Panel C), Task switching (Panel D). Sample includes users with college or more as described in text.

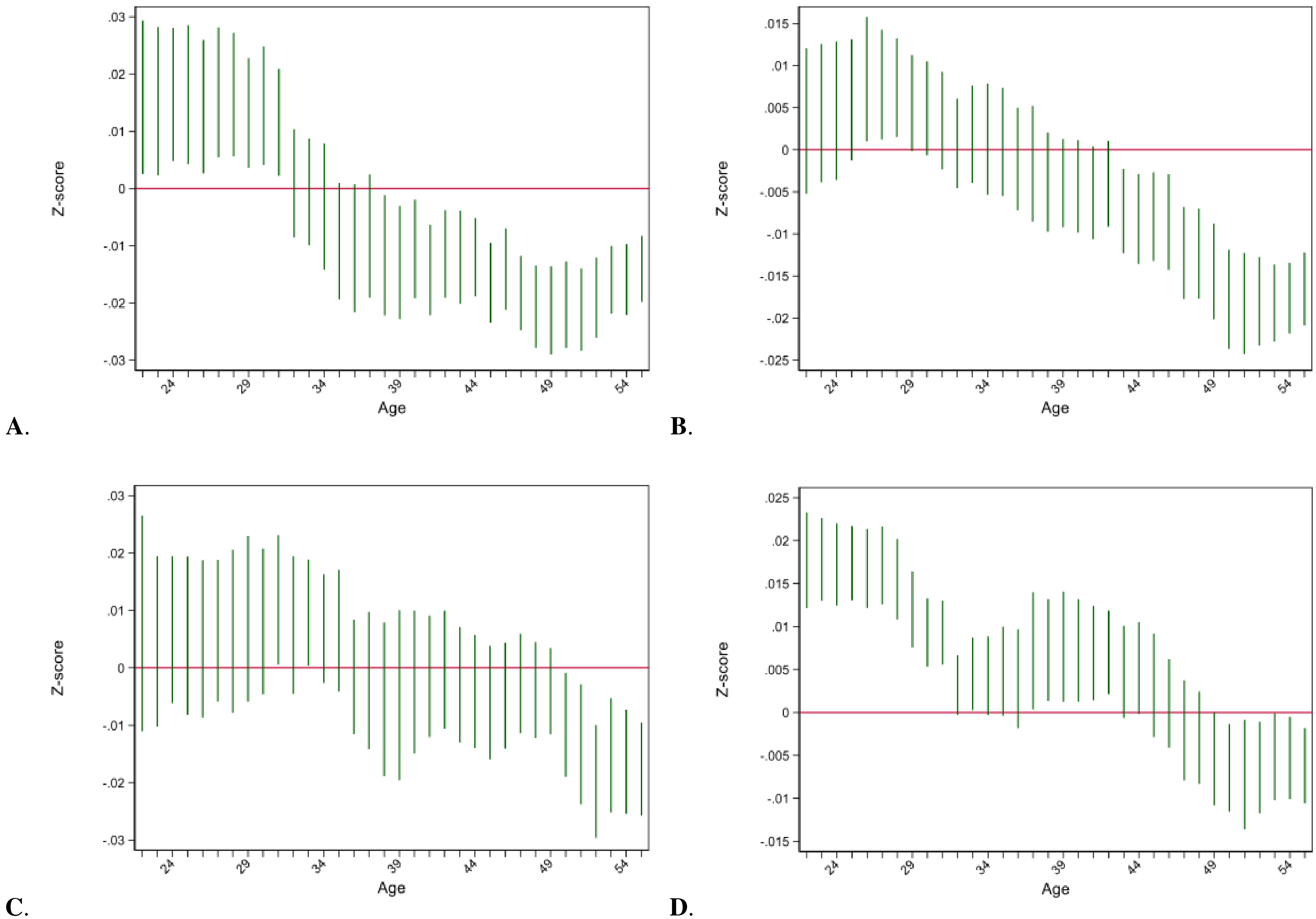


Figure 8: Age interacted with agglomeration across tasks
Notes: Numerical calculation (Panel A), Spatial recall (Panel B), Working memory (Panel C), Task switching (Panel D). Sample includes users with college or more as described in text. Agglomeration measures refer to individuals' ZIP3-level density.

Table 1: Descriptive statistics

	Numerical calculation [1]		Spatial recall [2]		Working memory [3]		Task switching [4]	
Panel A: Full sample								
A.1 Cognitive performance								
Average score (z-score)	0.30	(0.77)	0.14	(0.71)	0.09	(0.65)	0.16	(0.61)
Begginer score (z-score)	0.06	(0.76)	-0.13	(0.70)	-0.42	(0.61)	-0.39	(0.59)
Experienced score (z-score)	0.46	(0.82)	0.30	(0.78)	0.44	(0.73)	0.52	(0.70)
A.2 Circumstance								
Days to finish 20 sessions	205.2	(177.1)	168.8	(160.7)	327.3	(198.0)	179.6	(172.4)
Average mobile use in 20 sessions	0.47		0.54		0.37		0.64	
A.3 Schooling								
High-school or less	0.11		0.11		0.10		0.11	
Some college	0.23		0.24		0.23		0.24	
College+	0.66		0.65		0.67		0.64	
A.4 Demographics								
Male	0.43		0.46		0.42		0.44	
Age (years)	41.8	(11.5)	40.5	(11.5)	43.6	(11.2)	40.6	(11.4)
age 22 to 29	0.21		0.24		0.16		0.23	
age 30 to 44	0.33		0.35		0.30		0.35	
age 45 to 59	0.47		0.41		0.53		0.41	
Observations	184,511		242,749		49,359		243,640	
Panel B: College-plus sample								
B.1 Cognitive performance								
Average score (z-score)	0.47	(0.72)	0.22	(0.70)	0.17	(0.65)	0.23	(0.60)
Begginer score (z-score)	0.23	(0.71)	-0.05	(0.69)	-0.36	(0.61)	-0.32	(0.57)
Experienced score (z-score)	0.63	(0.77)	0.39	(0.77)	0.53	(0.73)	0.59	(0.68)
B.2 Circumstance								
Days to finish 20 sessions	203.3	(176.3)	167.7	(159.9)	328.4	(197.5)	178.7	(171.3)
Average mobile use in 20 sessions	0.46		0.53		0.36		0.62	
B.3 Demographics								
Male	0.43		0.45		0.43		0.43	
Age (years)	41.4	(11.3)	40.2	(11.3)	43.2	(11.2)	40.3	(11.2)
age 22 to 29	0.21		0.24		0.16		0.23	
age 30 to 44	0.35		0.37		0.33		0.37	
age 45 to 59	0.44		0.40		0.51		0.40	
Observations	121,801		157,698		33,122		156,337	

Notes: Standard-deviation in parentheses next to mean of continuous variable. In all samples 337 unique zip3 are included. Average scores are computed across 20 sessions while "begginer" score is average performance in the first 5 sessions and "experienced" score is average performance from the 16th to the 20th session. "Some college" group includes incomplete 4-year college attendees and community college graduates.

Table 2: Relationship between cognitive performance and agglomeration, College-plus sample

	Numerical calculation [1]		Spatial recall [2]		Working memory [3]		Task switching [4]	
Panel A: unconditional								
Zip3 density (in ln scale)	0.020		0.036		0.032		0.039	
	(0.005)		(0.004)		(0.006)		(0.003)	
CBSA Population (in ln scale)		0.005		0.013		0.014		0.025
		(0.006)		(0.008)		(0.008)		(0.007)
Sample	121,801		157,698		33,122		156,337	
Panel B: relative to numerical calculation performance								
Zip3 density (in ln scale)			0.017		0.025		0.024	
			(0.003)		(0.004)		(0.003)	
CBSA Population (in ln scale)				0.004		0.012		0.022
				(0.005)		(0.006)		(0.004)
Sample	27,109		27,109		27,109		27,109	
Panel C: conditional on demographics								
Zip3 density (in ln scale)	-0.003		-0.003		-0.003		0.006	
	(0.004)		(0.003)		(0.004)		(0.002)	
CBSA Population (in ln scale)		-0.010		-0.012		-0.011		0.003
		(0.004)		(0.003)		(0.004)		(0.003)
Sample	121,801		157,698		33,122		156,337	
Gelbach(2016) decomposition of omission bias on unconditional estimates								
Age	0.87	0.86	0.96	0.95	0.89	0.82	0.95	0.93
Post-grad education	0.03	0.02	0.00	0.00	0.01	0.01	0.00	0.00
Gender	0.14	0.24	0.05	0.09	0.06	0.10	0.04	0.06
Circumstance	-0.04	-0.11	-0.02	-0.04	0.04	0.07	0.01	0.01

Notes: Robust standard errors in parentheses clustered at zip3.

Table 3: Difference in agglomeration relationship: under 45 vs. 45+, College+ sample

	Numerical calculation [1]		Spatial recall [2]		Working memory [3]		Task switching [4]	
Panel A: Unconditional interaction								
Zip3 denslity (in ln scale)	0.023		0.017		0.017		0.015	
	(0.005)		(0.002)		(0.005)		(0.002)	
CBSA Population (in ln scale)		0.017		0.011		0.012		0.010
		(0.006)		(0.004)		(0.007)		(0.003)
Panel B: Interaction conditional on demographics								
Zip3 denslity (in ln scale)	0.019		0.014		0.014		0.013	
	(0.005)		(0.002)		(0.005)		(0.002)	
CBSA Population (in ln scale)		0.015		0.009		0.011		0.009
		(0.006)		(0.004)		(0.007)		(0.003)

Notes: Robust standard errors in parentheses clustered at zip3.

Table 4: Conditional relationship between gains in cognitive performance and agglomeration, College+ sample

	Numerical calculation [1]	Spatial recall [2]	Working memory [3]	Task switching [4]
Panel A: Unconditional				
Zip3 denslity (in ln scale)	0.001	0.001	0.002	0.011
	(0.001)	(0.001)	(0.003)	(0.001)
CBSA Population (in ln scale)	0.000	-0.002	0.002	0.014
	(0.001)	(0.001)	(0.003)	(0.002)
Panel B: Conditional on demographics				
Zip3 denslity (in ln scale)	0.001	0.002	0.002	0.011
	(0.001)	(0.001)	(0.003)	(0.001)
CBSA Population (in ln scale)	0.001	-0.001	0.002	0.013
	(0.001)	(0.001)	(0.003)	(0.002)

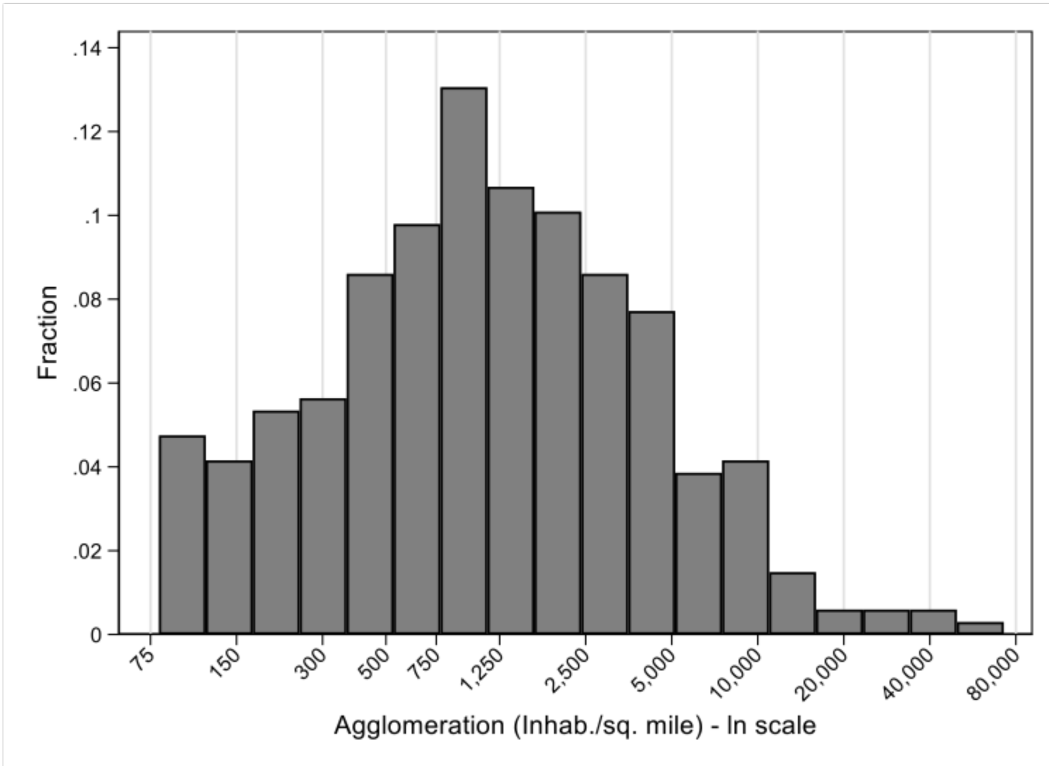
Notes: Robust standard errors in parentheses clustered at zip3.

Table 5: Difference in gains and agglomeration relationship: under 45 vs. 45+, College+ sample

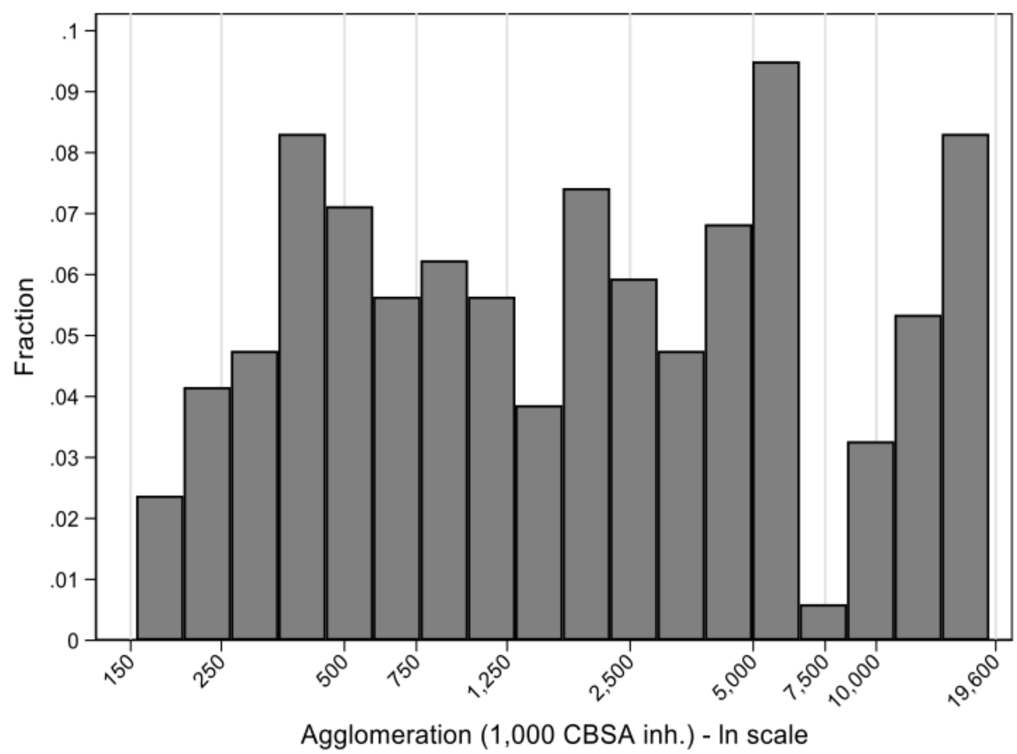
	Numerical calculation [1]	Spatial recall [2]	Working memory [3]	Task switching [4]
Panel A: Unconditional interaction				
Zip3 denslity (in ln scale)	0.003	0.004	0.007	0.008
	(0.002)	(0.002)	(0.005)	(0.002)
CBSA Population (in ln scale)	0.003	0.003	0.004	0.007
	(0.002)	(0.002)	(0.006)	(0.002)
Panel B: Interaction conditional on demographics				
Zip3 denslity (in ln scale)	0.002	0.002	0.008	0.009
	(0.002)	(0.002)	(0.005)	(0.002)
CBSA Population (in ln scale)	0.002	0.000	0.004	0.008
	(0.002)	(0.002)	(0.006)	(0.003)

Appendix Figures and Tables

Bacolod et al. (2020) “Learners in Cities: Agglomeration and the Spatial Division of Cognition”



A.



B.

Figure A1: Geographic representation of sample using 2 measures of agglomeration

Notes: Observations are one per ZIP3 with at least ten individuals engaging in at least one of four tasks from the cognitive-training platform described in the text. Agglomeration measures are based on individuals' ZIP3-level density and CBSA population computed from 2010 Census of Population data.

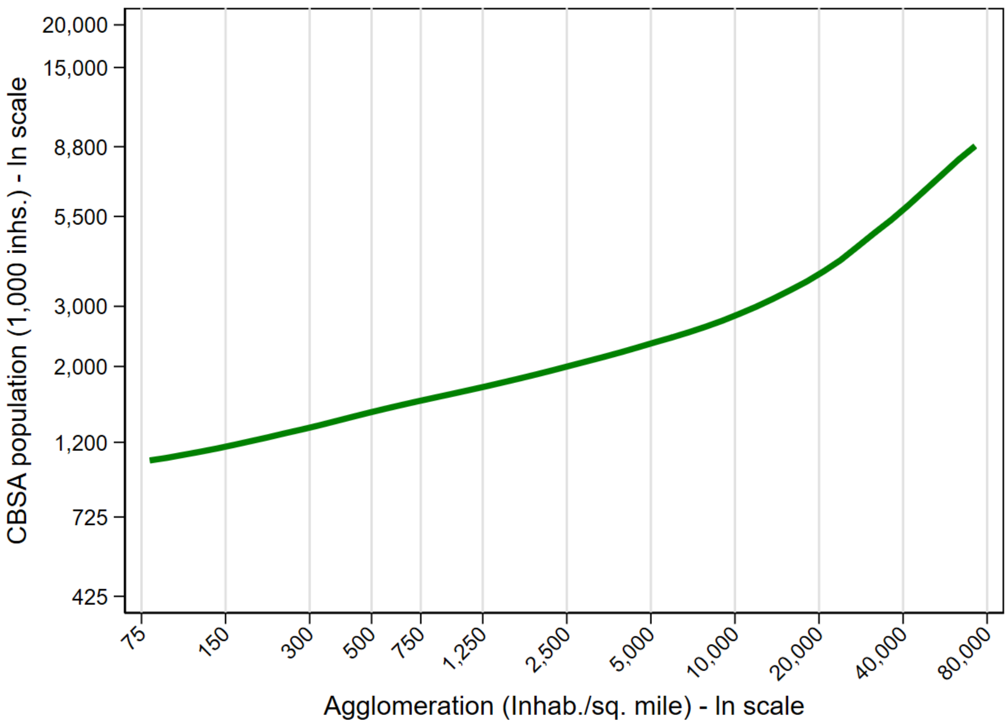


Figure A2: Relation between two measures of agglomeration

Notes: Observations are one per ZIP3 with at least ten individuals engaging in at least one of four tasks from the cognitive-training platform described in the text. Agglomeration measures are based on individuals' ZIP3-level density and CBSA population computed from 2010 Census of Population data.

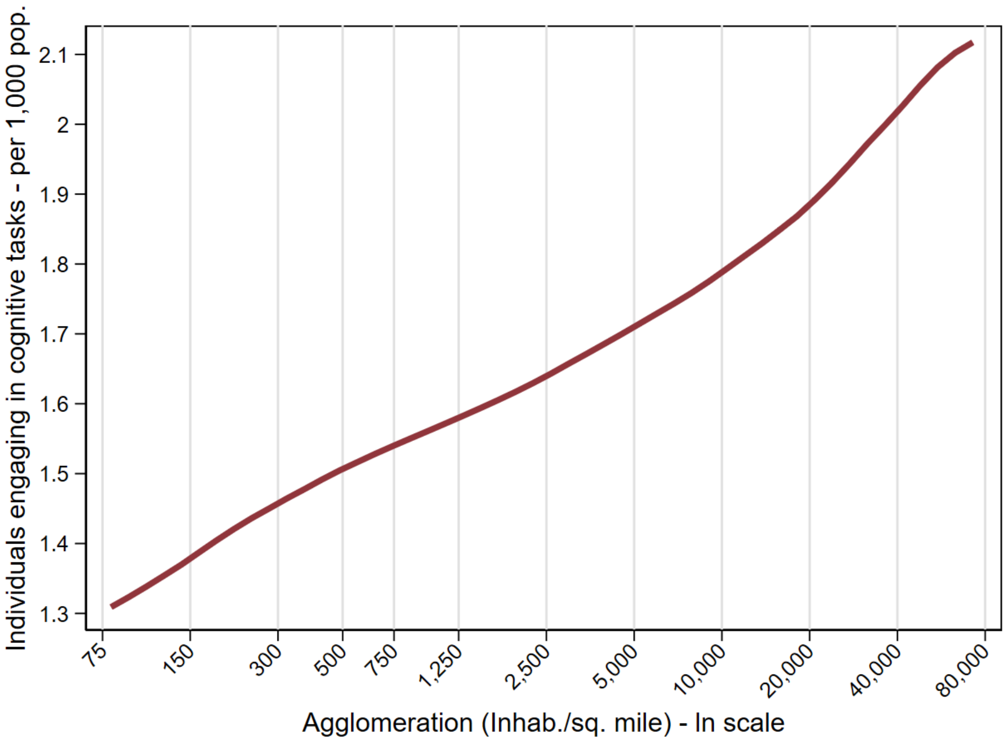
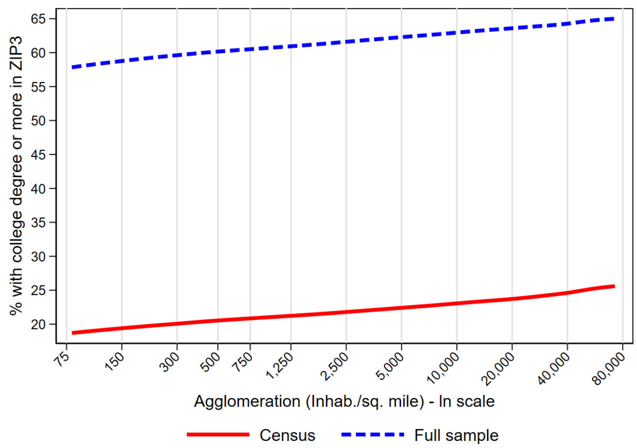
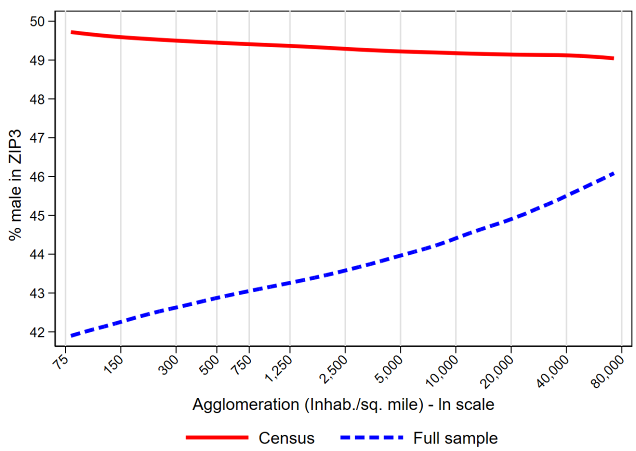


Figure A3: Relation between enrollment in cognitive-training (per 1,000 inhabitants) and ZIP3 density

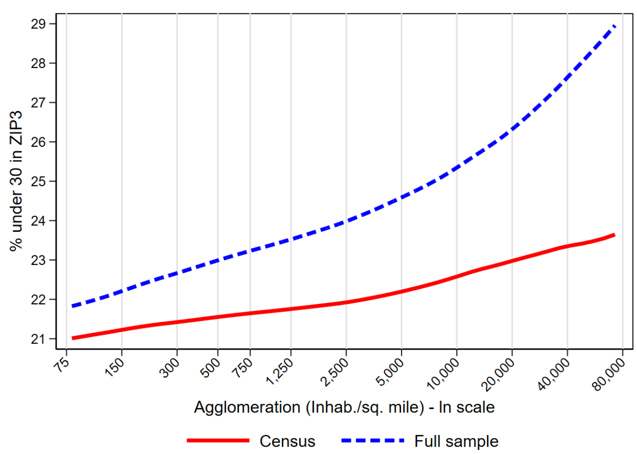
Notes: Observations are one per ZIP3. Count of is based on individuals (aged 22 to 55) engaging in at least one of four tasks from the cognitive-training platform described in the text. Agglomeration measure based on ZIP3-level density computed from 2010 Census of Population data.



A.



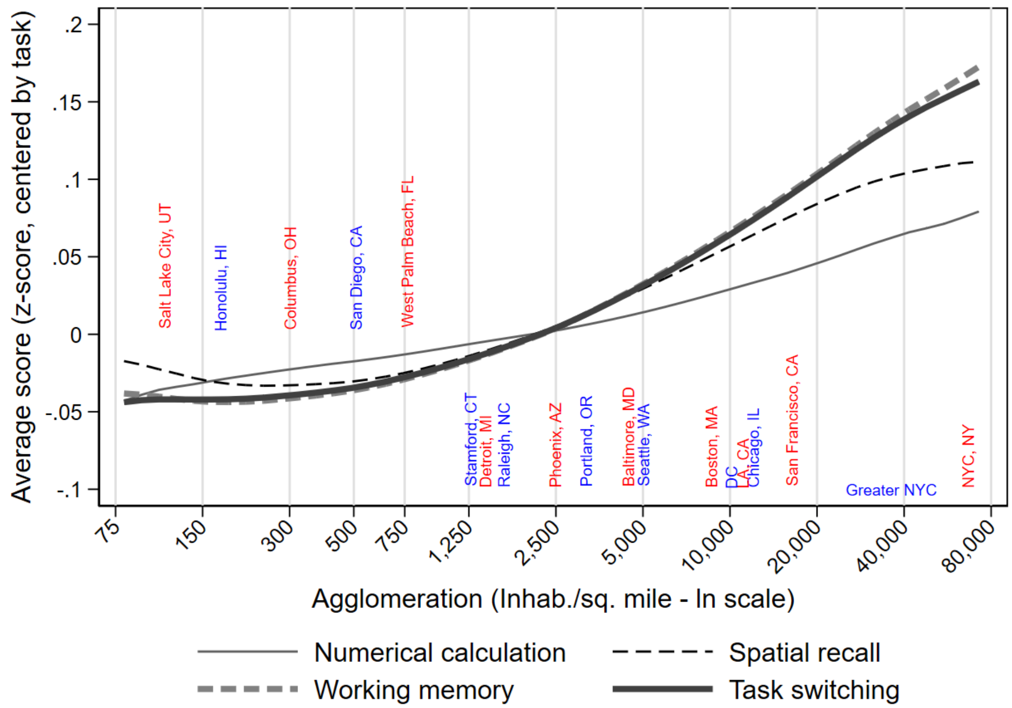
B.



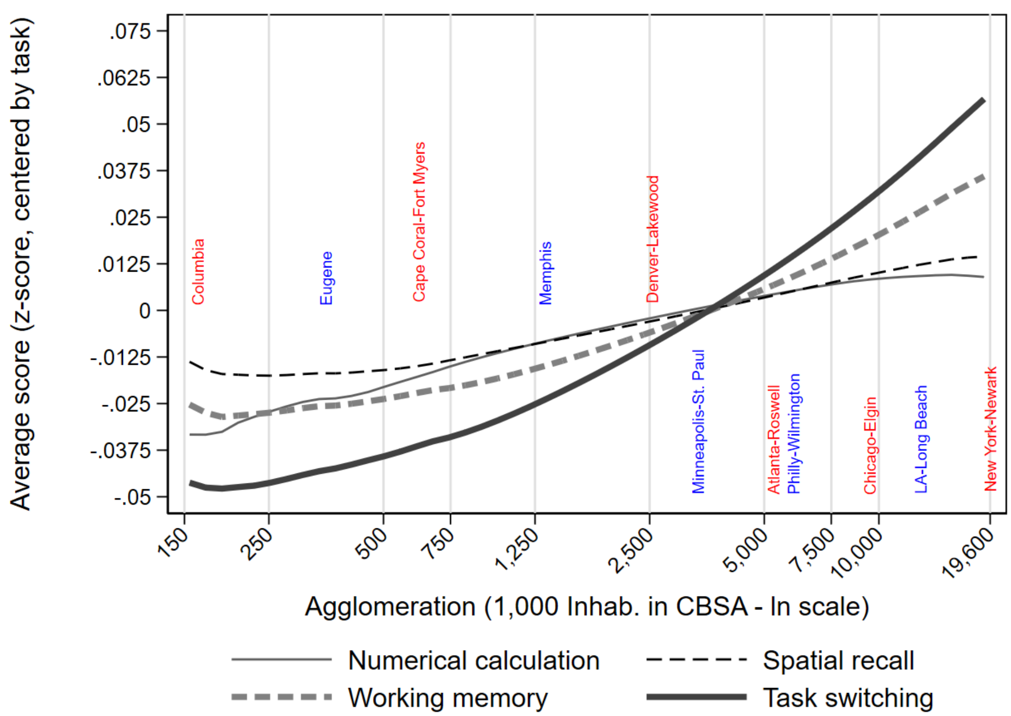
C.

Figure A4: Composition of full sample and of 2010 population

Notes: Observations are one per ZIP3. Count of is based on individuals (aged 22 to 55) engaging in at least one of four tasks from the cognitive-training platform described in the text. Agglomeration measure based on ZIP3-level density computed from 2010 Census of Population data.



A.



B.

Figure A5: Average performance in tasks and agglomeration of College+ sample who completed all 4 tasks

Notes: Sample restricted to college or more who completed all 4 tasks (n=27,109). Agglomeration measures refer to individuals' CBSA population and ZIP3-level density from 2010 Census data.

Table A1: Relationship between cognitive performance and agglomeration, Non-college sample

	Numerical calculation [1]		Spatial recall [2]		Working memory [3]		Task switching [4]	
Zip3 density (in ln scale)	-0.006		0.011		0.012		0.015	
	(0.005)		(0.004)		(0.005)		(0.002)	
CBSA Population (in ln scale)		-0.017		-0.001		0.003		0.007
		(0.005)		(0.005)		(0.006)		(0.003)
Sample	62,710		85,051		16,237		87,303	

Notes: Robust standard errors in parentheses clustered at zip3.

Table A2: Descriptive statistics of College+ sample who completed all 4 tasks

	Common characteristics [1]		Numerical calculation [2]		Spatial recall [3]		Memory match [4]		Task switch [5]	
Average score (z-score)			0.53	(0.70)	0.15	(0.69)	0.21	(0.63)	0.21	(0.58)
Begginer score (z-score)			0.28	(0.70)	-0.15	(0.69)	-0.34	(0.60)	-0.36	(0.56)
Experienced score (z-score)			0.71	(0.74)	0.34	(0.76)	0.58	(0.71)	0.59	(0.64)
Days to finish 20 sessions			149.8	(132.8)	125.5	(116.9)	326.9	(191.5)	137.7	(125.4)
Averege mobile use 20 sessions			0.32		0.34		0.32		0.36	
Age (years)	43.2	(11.2)								
age 22 to 29	0.16									
age 30 to 44	0.32									
age 45 to 59	0.51									
Male	0.44									
Observations	27,109									

Notes: See notes in Table 1.