

TECHNICAL APPENDIX: HOW EI WORKS AND HOW BARRETO APPLIES IT

Voting by secret ballot complicates any attempt to assess the demographic implications of legislative districts. We know how a locale voted, and we know the racial/ethnic makeup of the place, but we do not know the cross-tabulation between those two things; we cannot measure directly how voting differed by race and ethnicity.

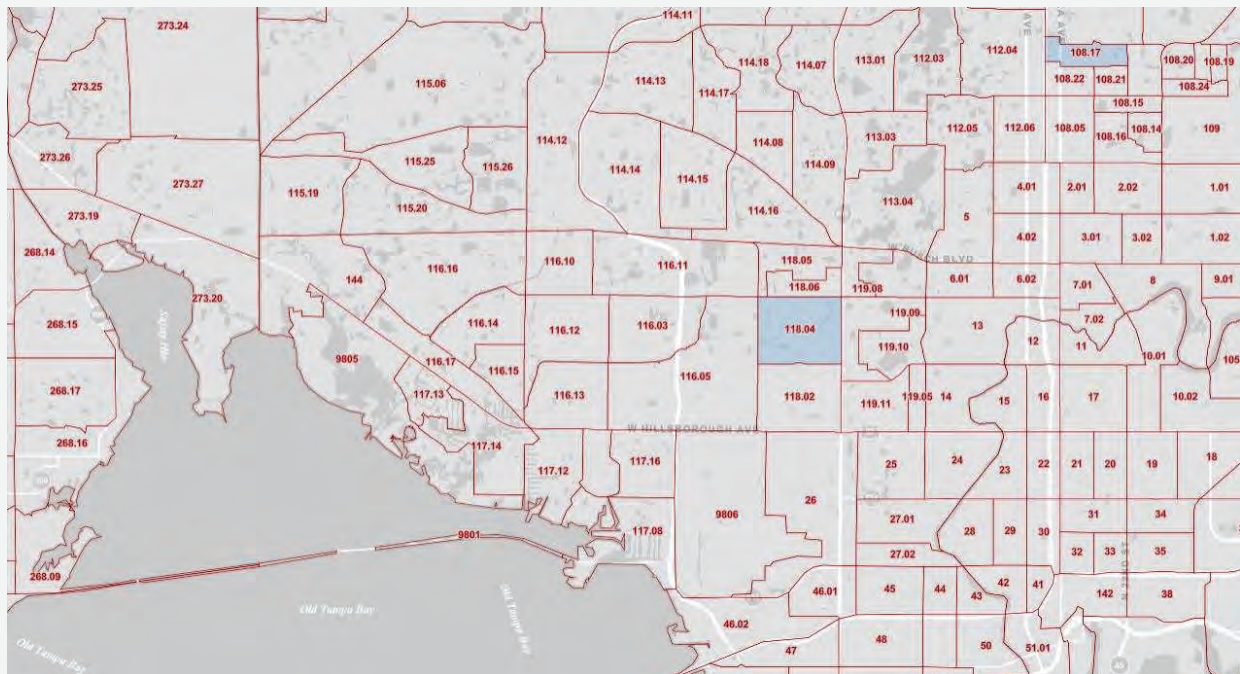
In some places, that ignorance is broad. Several steps in the electoral process might be hidden: racial/ethnic differences in voter registration, racial/ethnic differences in whether registered voters showed up to cast a ballot, not to mention what happens with those ballots. In especially data-rich environments like Florida, on the other hand, we know the race/ethnicity and the party registration of those going into the voting booth.

Still, we cannot follow voters into the booth, and we have no way to aggregate their choices by race and ethnicity. Even in the best of circumstances, therefore, an analyst is stuck trying to infer how race/ethnicity crosstabulates with vote choice – that is, how social groups differed in their voting behavior. We might know that 49.04% of SSD16 residents who voted in the 2018 Democratic primary were African American, and we might know that Gillum won 53.3% of the primary vote in that area, but we can only estimate what percentage of Black voters picked him versus picked one of his opponents, and we’re stuck guessing how everyone else voted as well. Notice the implication for voting-rights cases: The level of racially polarized voting – the gap between races in how they voted – can never be known factually. It can only be estimated using quantitative methods. Attempting to estimate those hidden quantities goes under the jargon “ecological inference.”

Gary King’s techniques for that purpose do not start out by estimating what’s happening across the entire area (for example, the entire county) being analyzed. Instead, the method starts with smaller units, such as precincts or the units Barreto created, and takes advantage of inputs the analyst *knows to be true* – the population demographics (from the Census) and the election returns (from the state’s election authority) – to restrict what it can guess for each little unit.

I will illustrate using a pair of Census tracts in Hillsborough County: 118.04 and 118.17. Their locations, relative to the north of Tampa Bay, appear in in Figure A.

FIGURE A – Forcing Ecological Inferences to be Mathematically Possible: Two Sample Census Tracts



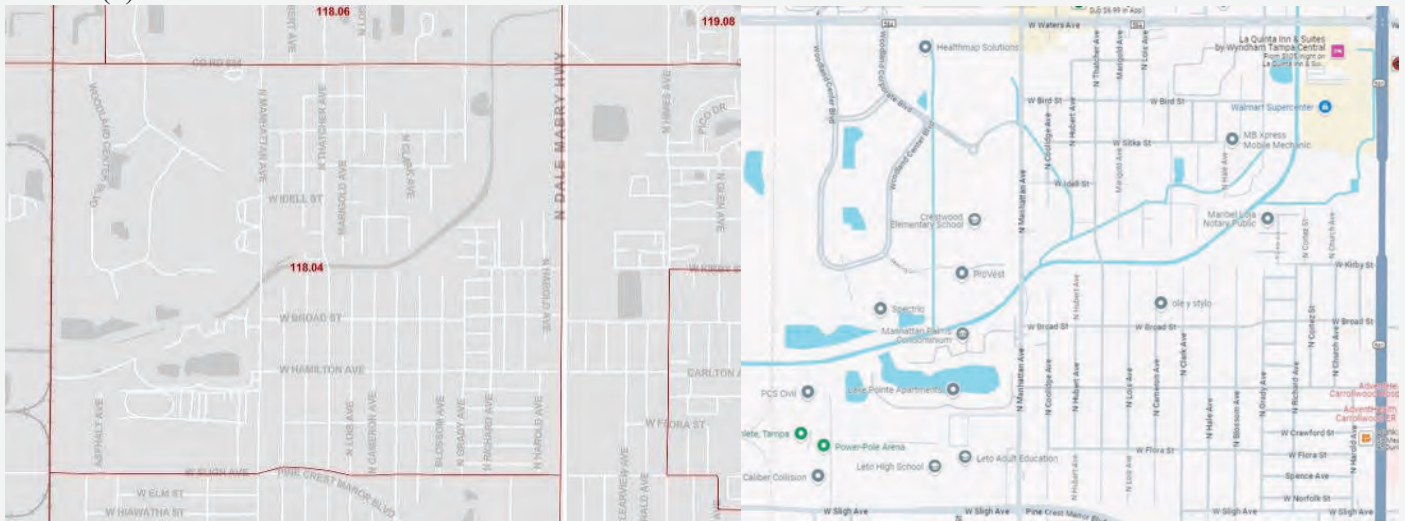
Hillsborough tract 118.04 appears as Figure B1, which has the Census Bureau map on the left and the Google map for the area on the right. It contained 4,564 adults at the time of the 2020 Census, of whom 3,524 were Hispanic. Most of that Hispanic population, more than $\frac{3}{4}$ of it, reported Cuban nationality. Perhaps for that reason, Hispanic voter registration only slightly tilts toward the Democratic Party in the tract. The political data accompanying Barreto's report suggest that the tract's Cuban residents were not as mobilized in 2020 as others living there; only 1,191 of the 2,045 people who turned out to vote apparently were Hispanic. The tract backed Donald Trump's presidential candidacy 55% to 45%.

Table A presents the inputs that would go into ecological inference for this tract, expressed both as counts and as proportions, and shows as question marks the crosstabulations we might need to know: how Hispanics voted, and how everyone else did. Here's how King's method ensures that estimates will be mathematically possible for each of these little units, a process called the **method of bounds**:

Trump received 1,129 votes, but only 854 non-Hispanic voters turned out, so Trump's Hispanic support there could not have been 0%. At a minimum, he picked up $1,129 - 854 = 275$ Hispanic votes. That is, at least $275/1,191 = 23.1\%$ of Hispanic voters backed Trump. At the same time, more Hispanics showed up than Trump received votes in the tract, so Trump could not have received 100% of the Hispanic vote; at least $1,191 - 1,129 = 62$ Hispanic voters must have picked Biden, even if not a single non-Hispanic backed Trump.

FIGURE B – A Cuban Census Tract and a Puerto Rican One

(1) Tract 118.04



(2) Tract 108.17

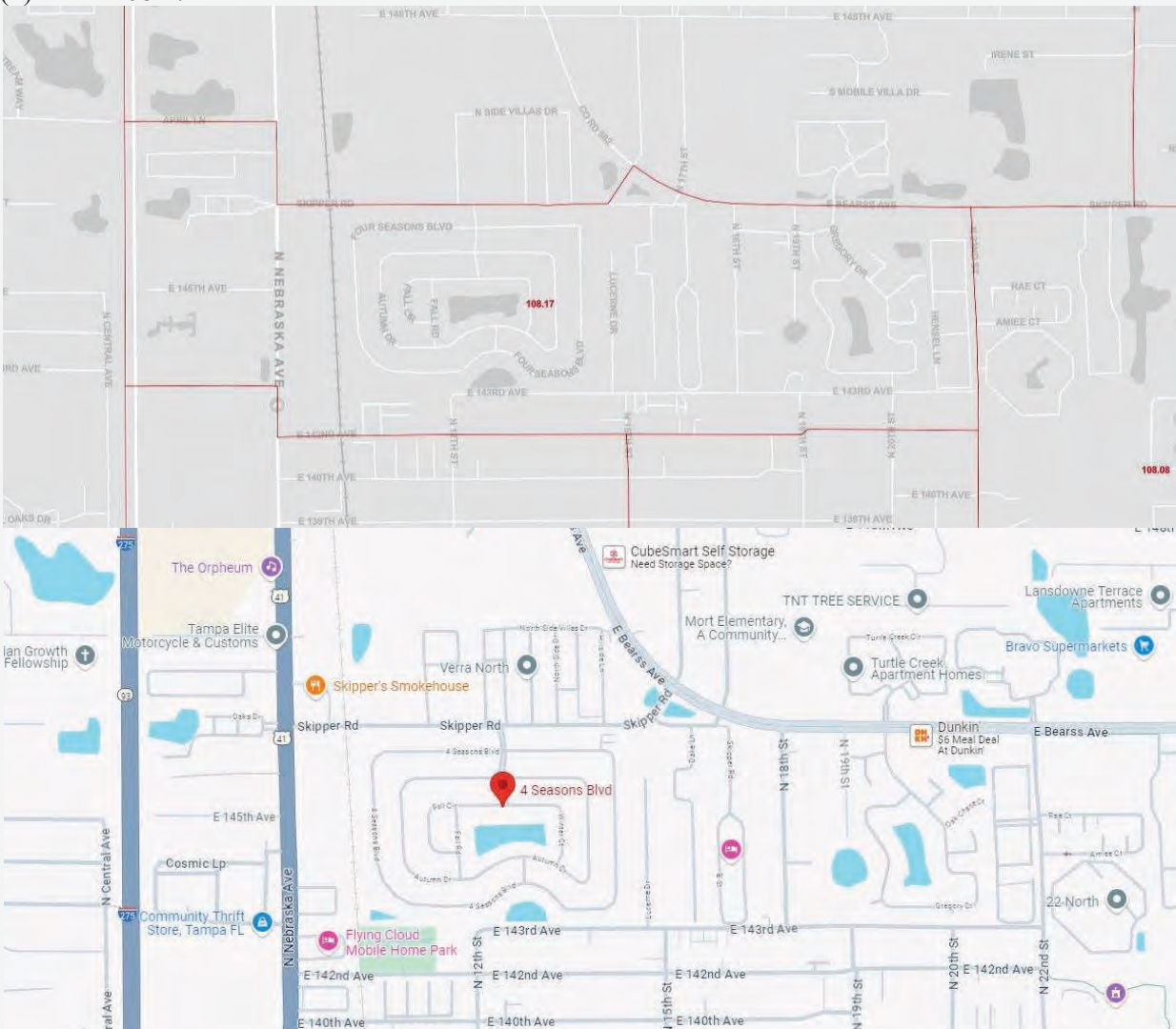


TABLE A – The Method of Bounds in a Heavily Cuban Precinct

Hillsborough 118.04

	Hispanic	Non-Hispanic	
Biden	?	?	916
Trump	?	?	1129
	1191	854	2045

Hillsborough 118.04

	Hispanic	Non-Hispanic	
Biden	?	?	0.45
Trump	?	?	0.55
	0.58	0.42	2045

We cannot narrow down, in the same way, how non-Hispanic voters behaved. Anything from 0% Trump support to 100% Trump support would be mathematically possible here. Yet we still know a lot about how those non-Hispanic voters could have behaved because once we know Trump's rate of Hispanic support, then one and only one rate of non-Hispanic support would be possible mathematically. That is, the support for Trump is linear:

$$\begin{aligned} \text{Trump Vote} = & \text{Hispanic Turnout} \quad \times \text{Rate of Hispanic Vote for Trump} \\ & + \quad \text{Non-Hispanic Turnout} \quad \times \text{Rate of non-Hispanic Vote for Trump} \end{aligned}$$

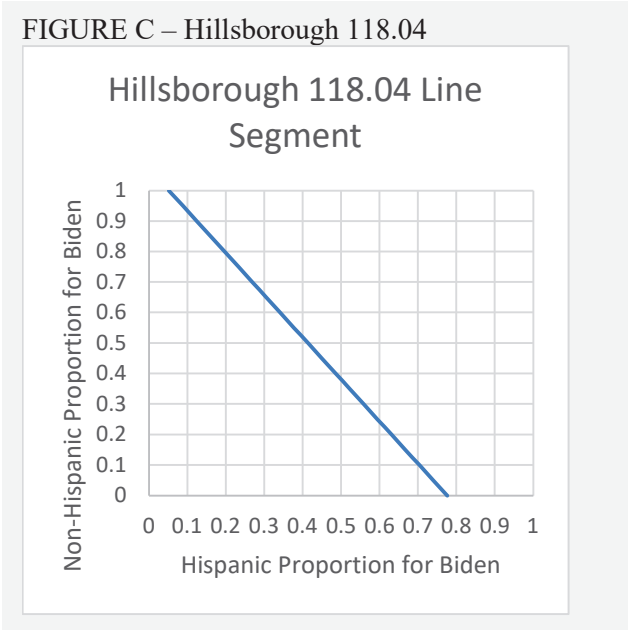
We know the two turnout rates in the tract, and once we hypothesize a particular level of Hispanic support, then the corresponding level of non-Hispanic support would be computed this way:

$$\text{Non-Hispanic Rate} = (\text{Trump Vote} - \text{Estimated Hispanic Votes for Trump}) / \text{Non-Hispanic Turnout}$$

Obviously neither of these rates can fall below 0% or go higher than 100%, so if we were going to graph what's possible for this particular Census tract, the result would be a line segment rather than a line. The line segment for Hillsborough tract 118.04 appears in Figure C, illustrating possible rates of support for Biden rather than Trump. The location of that line segment indicates what we've already determined from simple calculations: Because the line segment extends from top to bottom, the non-Hispanic rate of support for Biden can range from 0 – 100%, whereas the line segment does not extend from left to right – showing that Biden's Hispanic support could not have been greater than $100 - 23.1 = 77.9\%$ (because that's the farthest right that the segment reaches), but also was not zero (because the left-hand side of the segment never reaches the left-hand side of the box).

Because what's possible for each group depends on the size of that group in the unit's population, with our certainty about how the group voted depending on the relative

size of the group, the slope of the line segment also tells us which racial/ethnic group is most numerous in the locale. A line that is either vertical or horizontal is homogenous; we know precincts how one group voted but have no idea about the other group. A locale that’s almost equally balanced between the two groups, as Hillsborough 118.04 was balanced between Hispanics and non-Hispanics, will cut diagonally across the box, because either group could have given high or low support to the candidate. We know a lot less about tracts such as this one. But the true combination of Hispanic and non-Hispanic support for Biden appears somewhere on that line segment, and when King’s method tries to estimate what those rates were, it will only pick a spot somewhere along that segment.



The second sample Census tract, 108.17, appears as Figure B2. This tract also contains a large Hispanic population, but of Puerto Rican rather than Cuban descent – a group with stronger ties to the Democratic Party. Table B shows why we will have a much easier time estimating political behavior in this tract, compared to the last. Joe Biden received 798 votes here (or 83%). Even if every Hispanic cast a vote for Biden, at least $798 - 416 = 382$ non-Hispanics (or 70.3% of them) must have sided with him. Even if every non-Hispanic backed Biden, at least $798 - 543 = 255$ (or 61.3%) of Hispanics must have backed him. So even though the line segment for this tract will be angled about like the last one, because the population is fairly evenly balanced like in the last one, we’re still going to be

TABLE B – The Method of Bounds in a Heavily Puerto Rican Precinct

Hillsborough 108.17			
	Hispanic	Non-Hispanic	
Biden	?	?	798
Trump	?	?	242
	416	543	959

Hillsborough 108.17			
	Hispanic	Non-Hispanic	
Biden	?	?	0.83
Trump	?	?	0.25
	0.43	0.57	959

able to narrow what's mathematically possible to a much greater extent here because of the very high level of Biden support. Biden did so well that both groups had to support him at high levels. Figure D shows the line segment that captures all of the possible combinations of Hispanic and non-Hispanic support for Biden that would be mathematically possible in this tract. And we've done all of this narrowing down without estimating, guessing, or assuming a single thing so far.

If we combine all the line segments for every single tract in an analysis into a single box – that is, if we collect everything that's mathematically possible for each tract in one place – we get what King calls a **tomography plot**. An experienced EI user, who has looked at a lot of tomography plots and analyzed a large variety of datasets, can tell a lot about whether ecological inference is likely to work – and what problems might plague it

FIGURE C – Hillsborough 108.17

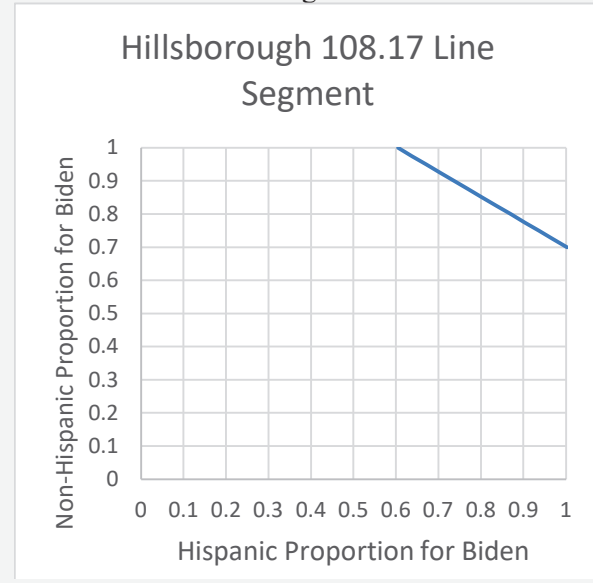
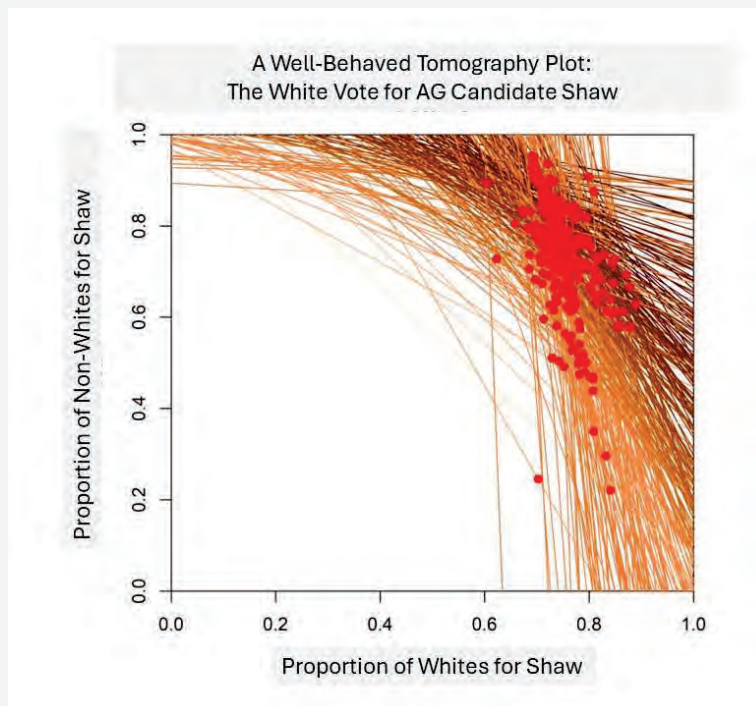


FIGURE D –



NOTE: The horizontal, lateral, and vertical lines all tend to converge around the same spot in the upper-right-hand corner of this tomography plot. For that reason, it is fairly easy to identify the region of the square where the combination of White and non-White candidate support is most likely to appear.

– from the visualization of all those line segments.

For example, FIGURE D shows a tomography plot for White voting behavior at the tract-level, when using racial/ethnic turnout data, for the 2018 Democratic primary contest between Shaw (the Black candidate of choice) and Torrens. Each line segment represents one tract in Hillsborough County, with each tract's true combination of White and non-White support for Shaw appearing somewhere on the line segment associated with that tract. In this case, because the line segments – horizontal, lateral, and vertical – all appear to be passing through roughly the same section of the unit square, EI will not have a hard time inferring support rates. The red dots represent the best guesses for each line segment, with most of them clustered where the line segments are coming together.

Once King's method picks a spot on each of these line segments, with each spot representing a mathematically possible level of combined White and non-White support for Shaw, those tract numbers are added up to represent an estimate for how Whites and non-Whites voted in the whole county. Because the county-level estimate builds from a whole series of tract-level estimates that are mathematically possible, the method's guess for what happened in the county also will be mathematically possible. Indeed, because each step has been disciplined by the method of bounds – because all the guesses are bounded by what each tract tells us about what is and is not possible – then the method sometimes can perform well even when some of its underlying assumptions are not met.

That voting in a smaller unit can be reduced to a line segment is only one assumption that King's method makes. King's method also makes a simplifying assumption to help with deciding where to on the line segment a unit's voting rates are likely to appear. He assumes that each group has the same basic underlying political preferences everywhere being analyzed, give or take the usual randomness in human behavior. African Americans will have some typical level of support for Gillum, although it might pop higher or lower randomly in any given precinct. The same can be said for Hispanics and White/Other voters: They have a basic lean, but it'll jiggle a bit from precinct to precinct. That is why it was important, in the last tomography plot, to see the various line segments converging in roughly the same location. Where that spot appeared along the two axes would be EI's guess for the two underlying rates of group behavior. Trying to make ecological inferences when that assumption is faulty, though, can blow up the estimation unless the method of bounds forces the estimates back to reality or unless the analyst actively captures the diversity within the group being studied.

King's method can be even more vulnerable to error if the differences in a group's voting behavior are not random. As mentioned earlier, King's method derives estimates for the larger unit (such as the county) using the information available in smaller units (such as tracts or precincts). Those smaller units might be

fairly diverse. Some might be overwhelmingly Black, others overwhelmingly White, others overwhelmingly Hispanic. The vote percentages might vary across those places as well, with some units heavily supporting one candidate while others heavily support a different candidate. By looking at how the vote choice varies with a place's demographic composition, we can guess how different groups might be voting. If we notice that Gillum's support tends to grow as the little units have higher Black density, for example, then we'll infer that African-American voters tended to back Gillum. A third assumption King's method makes is that voting by one group does not depend, directly or indirectly, on the size of the other group. The patterns EI uses to estimate voting rates can be misleading, though, if how a group behaves depends on who lives around them. If Whites living in heavily Black neighborhoods vote differently than Whites in heavily White areas, then that can throw the estimation.

Naïve ecological inference, conducted with no regard for such contextual patterns, can introduce "aggregation bias" into the estimation – not just getting the results wrong in certain places, but getting the wrong results for the whole area being studied. In particular, if Whites or Hispanics show greater Democratic support in mixed-race areas than they do in homogenous ones – for example, if urban Hispanics and Anglos are more Democratic than rural and small-town ones – the bias will be toward making racial polarization seem greater than it actually is.

Armed with this deeper understanding of how King's solution to the ecological-inference problem works, it is now possible to explain in more technical terms why Barreto's analysis would go awry. Barreto runs up against each of the three potential problems implied just now. Barreto does not employ ecological inference in a reliable fashion consistent with best practices, and each of the three difficulties faced by his analysis can be expressed in terms of tomography plots.

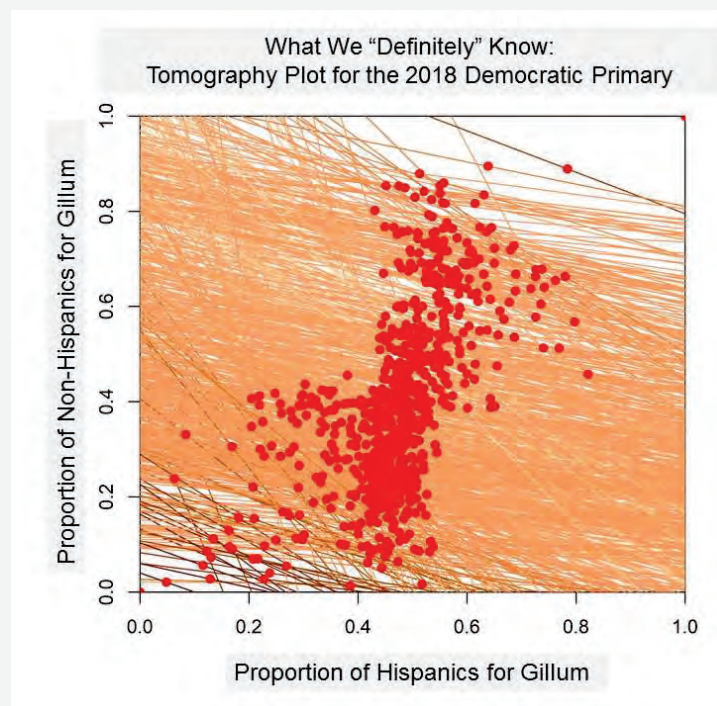
First, although in most instances of ecological inference, a low-level unit could be represented by a line segment, the tomography plot wouldn't capture what's possible for Barreto's analysis. He's using Census blocks partitioned in multiple ways and then grouped: first by county, then by Census place (similar to a "municipality"), then by Census tract, then by voting-tabulation district (VTD, similar to a precinct), and finally divided up by the enacted State Senate district where it falls. This complex unit of geography has no name or intuitive meaning, so I've just been calling them "Barreto units." What's troublesome about these Barreto units is that we don't actually know the political behavior of the Census blocks used to create them. Votes aren't counted by Census block. All we have are guesses as to how those blocks voted, created by projecting election returns down to that level. And we have no way to know how messy those projections are. So when Barreto's ecological inference attempts to use the method of bounds to restrict his estimates to what's mathematically possible, he's not actually feeding things we "definitely know" into the system. He's feeding estimates with an

unknown amount of error into the system. The only reason those Barreto units can show up in the tomography plot as line fragments is because he's ignoring that error and treating the numbers as truth. Not only will the resulting ecological inferences be less accurate, because of the measurement error, but the software will attach greater certainty to its ecological inferences because it doesn't know it was fed bad data.

Now the second assumption. If it's true, as assumed, that the two groups in an ecological inference each have a common underlying level of support level for the candidate across the entire county, then the line segments should tend to crisscross near the spot representing those two values, the way they did in Figure D. Instead, as seen in Figure E, the tomography lines for Hispanic voting in the 2018 gubernatorial primary do not converge on a particular space in the unit square. The analysis is combining heavily non-Hispanic units that are mostly White – they will tend to be the horizontal lines near the bottom of the square – with heavily non-Hispanic locations that likely are heavily Black and appear as horizontal lines near the top of the square. Non-Hispanics do not at all vote in a similar way across the county!

Meanwhile, the lack of vertical or even especially slanted line segments in Figure E tells us that the data do not provide much information about Hispanic voting. What we can see from the mass of red dots in the center is that EI, assuming that all Hispanics tend to vote the same way everywhere, keeps offering Gillum support of around 45% as the best guess for each unit. For the Barreto units

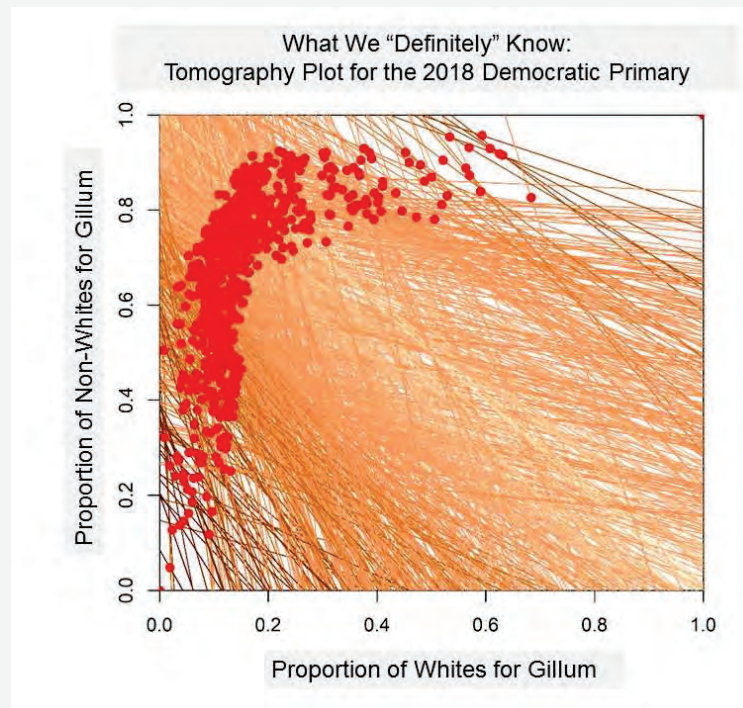
FIGURE E –



where we don't know a lot about Hispanic voters because they're small in number, ecological inference keeps guessing roughly the same thing. Yet Florida's Hispanic population includes both Cubans and other Hispanics (including Puerto Ricans, Mexicans, and Venezuelans). No one seriously would propose that those groups all exhibit the same underlying political behavior. And we can see from the tomography plot that actual Hispanic behavior is much more varied than EI generally assumes. Why? Because in the few places where the method of bounds forces estimates away from the middle, we see two clusters: one on the bottom left, and another toward on the top right. Because the assumptions of the model are wrong, we're likely to see instability in estimates of Hispanic voting, which in turn will mean that estimates for Black and White voting will be less accurate than necessary to guess how district lines actually will perform. And indeed, estimates of Hispanic support for Gillum varied so widely that I didn't even bother to include them in Table 5. They ranged from heavy support for the Black candidate of choice all the way to negligible support for that candidate.

Now, it's true that both versions of EI that Barreto uses – Iterative EI and RxC EI – would be estimating separately how Blacks and Whites voted. To the extent it can get those numbers right, the analysis could adjust for the messiness of the Hispanic analysis. But Figure F shows another sign that the ecological inferences are likely to be sloppy, one that might contaminate White and Black estimates as well. Figure F offers a tomography plot from the same 2018 Democratic primary, but this time focused on White voters. Once again, the line segments are not mostly crisscrossing in one spot. And once again, the red dots

FIGURE F –



show that how a group votes likely depends on the size of the other group. The way the red dots sweep up and then rightward tells me that White voters differ by the racial context of their neighborhood, the sort of pattern that can cause aggregation bias unless taken into account. As we move upward in the square, likely moving toward units with a heavier African-American presence, it appears as though White support for Gillum increases as well, because the red dots start drifting rightward. To some degree, ecological inference likely is attributing that changing White behavior to their Black neighbors. That's the aggregation bias.

A final note. To test what happens when we use low-level units to estimate voting behavior, but then aggregate up to the Senate district level, we tried that approach using block groups and tracts. The rows added to Table 5 here show that the results aggregated up to those levels are fairly close to what EI returned when directly estimating the vote at those higher levels. The aggregation process works fairly well.

TABLE 5 – Racial Polarization Depending on Estimation Method

	HILLSBOROUGH						PINELLAS					
	Whites			Blacks			Whites			Blacks		
	Graham	Gillum	TOTAL	Graham	Gillum	TOTAL	Graham	Gillum	TOTAL	Graham	Gillum	TOTAL
Barreto Code												
Barreto RxC	58.1	15.2	95.3%	8.1	86.8	95.6%	57.1	11.3	93.4%	8.6	75.6	103.1%
	55.1	22.4	n/a	3.2	88.8	n/a	55.4	15.2	n/a	5.2	74.0	n/a
Block Groups												
Tracts	49.7	23.6	98.0%	4.2	89.1	105.8%	51.6	14.8	99.3%	6.7	75.0	104.2%
	50.7	21.0	97.3%	1.3	91.0	103.5%	52.1	14.0	99.2%	4.8	79.7	106.4%
Barreto --> BG's												
Barreto --> Tracts	52.1	24.2	99.2%	4.4	91.2	102.7%	55.2	15.4	100.0%	5.3	76.4	101.6%
	53.2	21.1	97.9%	1.4	92.6	101.7%	55.5	14.5	99.3%	3.6	82.0	104.3%
Barreto Units												
Block Groups	54.5	20.3	97.9%	6.9	82.3	95.8%	52.3	18.6	99.7%	9.7	70.5	100.4%
Tracts	51.5	21.3	99.4%	8.2	74.9	101.6%	50.2	17.0	99.6%	10.2	67.1	101.6%
	52.3	19.0	98.8%	6.8	77.6	101.7%	50.5	16.2	99.3%	8.6	70.6	103.1%

NOTE: Iterative EI shows high instability depending on the unit of analysis used. The method improves the estimation of White and Black voting when using turnout data rather than voting-age population. TOTAL represents the summed support for all candidates; deviating significantly from 100% is a sign of estimation trouble.

1

https://www.researchgate.net/publication/376798358_Familiarity_Doesn't_Breed_Contempt_The_Political_Geography_of_Racial_Polarization

² Voss, D. Stephen. 1996. "Beyond Racial Threat: Failure of an Old Hypothesis in the New South." *Journal of Politics* 58:1156-70.

³ For example, Lublin, David, and D. Stephen Voss. 2000. "Racial Redistricting and Realignment in Southern State Legislatures." *American Journal of Political Science* 44(October):792-810.

⁴ King, Gary. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton, NJ: Princeton University Press. Pp. xxii, 22-24.

⁵ Voss, D. Stephen. 2004. "Using Ecological Inference for Contextual Research: When Aggregation Bias Is the Solution as Well as the Problem." In Gary King, Ori Rosen, and Martin Tanner (eds.), *Ecological Inference: New Methodological Strategies*. New York: Cambridge University Press. Pp. 69-96.

6

https://www.researchgate.net/publication/376798358_Familiarity_Doesn't_Breed_Contempt_The_Political_Geography_of_Racial_Polarization

⁷ Voss, D. Stephen, and David Lublin. 2001. "Black Incumbents, White Districts: An Appraisal of the 1996 Congressional Elections." *American Politics Research* 29(March): 141-82; Voss, D. Stephen, and Penny Miller. 2001. "Following a False Trail: The Hunt for White Backlash in Kentucky's 1996 Desegregation Vote." *State Politics and Policy Quarterly* 1(March):63-82; Lublin, David, and D. Stephen Voss. 2002. "Context and Francophone Support for Sovereignty: An Ecological Analysis." *Canadian Journal of Political Science* 35(March):75-101;

⁸ Voss, D. Stephen and David Lublin. 1998. "Ecological Inference and the Comparative Method." APSA-CP: Newsletter of the APSA Organized Section in Comparative Politics 9(1):25-31; Voss, D. Stephen, and Penny Miller. 2017. "The Phantom Segregationists: Kentucky's 1996 Desegregation Amendment and the Limits of Direct Democracy." *Commonwealth Review of Political Science* 4(1): 21-38.

9

https://www.researchgate.net/publication/268341180_Racial_Polarization_and_Turnout_in_Louisiana_New_Insights_from_Aggregate_Data_Analysis

¹⁰ Computers cannot behave randomly, although they can imitate it. When analysts want a computer program to "randomly" generate numbers, but to do so in a way that can be repeated later or by someone else, they set a "seed" ahead of time. The same analysis, performed later with the same seed, will produce exactly the same results. Barreto's code attempted to set a seed, but failed to do so because of a typing error. Not knowing what seed his computer used, no one would be able to duplicate his results exactly in order to

evaluate them.

¹¹ Strictly speaking, what I obtained were the block assignment files needed to construct these maps.

¹² SSD16 earned a Reock score of .38 and a Polsby-Popper score of .36, both better than the scores earned by more than half-a-dozen other districts statewide. The first indicates that SSD16 does not deviate excessively from a theoretical “perfect circle,” while the latter means that the district does not have a lot of jagged edges or tendrils relative to its overall size. Finally, the best compactness measure that DRA reports – the Know It When You See It (KIWYSI) score – puts SSD16 at a middling 50 out of 100, better than several other Florida Senate districts.

¹³ I say “as if” because, in practice, neither the original mapmakers nor Dr. McCartan tried to contain the region’s five (5) districts within the counties defining the MSA. They linked the MSA’s outlying territory with adjacent counties outside the MSA. But that’s not necessarily sinister. Counties at the edges of an MSA sometimes have their largest population, and their densest population, close to and economically integrated with the central city, while the far-flung smaller communities elsewhere in the county may resemble the adjoining rural counties. Efforts to break apart an MSA need to be judged on a case-by-case basis.

¹⁴ I say “obviously” because former Green Beret Jay Collins, a Republican, defeated the incumbent left over from the old maps, Sen. Janet Cruz, by a 54.4%-45.2% margin. But note that the new SSD14 that she lost was not significantly different in partisan competition from what her old SSD18 had become. The district was 49.76% Democratic and 48.15% Republican under the benchmark map, compared to 49.9% D and 47.98% R now. It was an exceptionally good year in Florida for Republicans.

¹⁵ I make no judgment whether Ruskin or its specific geopolitical boundaries are more or less worthy of respect than other Florida places or Florida boundaries. I simply note that these changes do violence to both, despite the expectation that districting will respect these features of the map. Clearly other redistricting priorities have shaped the ACLU proposals.

¹⁶ I say “these days” because at one time, such cracking of the Black vote was less about partisan goals – often the mapmakers and the African-American voters were both Democratic constituencies – and more uniquely racial in intent.

¹⁷ Barreto claims in the text that he analyzed Pinellas County, but I have no evidence that claim is accurate. The code he provided sets up no such analysis (although applying it to Pinellas requires relatively trivial changes), and he furnishes no tables or figures from Pinellas to document the results. Finally, I note that in the part quoted here, he’s only referring to Hillsborough. So we were left guessing whether the reference to “Hillsborough and Pinellas County” is vestigial language, from a time when Barreto intended to perform a Pinellas County analysis that he didn’t actually get to complete, or if he actually did analyze Pinellas but excluded the results from his report.

¹⁸ Technically speaking, King’s method assumes that the group’s voting behavior across small units follows a bell-shaped curve and can be expressed, in conjunction with the comparison social group’s voting behavior, as following a bivariate normal distribution.

¹⁹ Counsel did not ask me to develop my own ecological inferences, only to assess and apply Barreto’s

approach. I am not, however, offering this criticism without a clear idea of what I would have attempted to improve the results. First, the Census Bureau collects and reports Hispanic nationality for the voting-age population at the tract level (although it's imperfect because they only report groups of 900 or larger). Second, the data Barreto used included information on the party registration of the Hispanic voters who cast ballots. Either of those data sources could have been incorporated into the ecological inferences performed here, to see whether they added stability to the estimation.

²⁰ He makes this claim in Appendix C on page 25, near the end of footnote 19: “we compared our results with models using VAP by race and ethnicity—for which a standalone non-Hispanic white race category exists in the data—and found substantively identical results.” I collapsed White and Other for VAP, so that the demographic categories would be the same for VAP and Turnout, and the results when estimating the vote in a Democratic primary are not what I would call substantively similar. The differences in what EI estimates for Hispanics are especially dramatic, which might not matter in some portions of Florida, but definitely can matter in Senate districts on the peninsula.

²¹ We exclude the Hispanic results partly because they're so unstable, and partly just so the table would fit.

²² Judging from their instructions, that appears to be how the data's creators envisioned their block-level estimates being used.