Re: Item No. 2(ii)



Estimation and Projection of HRM Possible Voters at the DA Level – 2009 and 2012

Tony Lea, Ph.D., SVP, Chief Methodologist Rupen Seoni, M.A., VP, Practice Leader

April 2010

Agenda

- ▲ Environics Analytics who are we
- Project Objectives
- ▲ Data Inputs
- Overview of Methodology
- ▲ Methodology (A)- Large Geography Scale
- ▲ Methodology (B) Small Scale 2009
- ▲ Methodology (B) Small Scale 2012
- ▲ Final Aggregation to the HRM Electoral Districts
- ▲ District Level Estimates
- ▲ Questions



Organization: Environics Group of Companies



A Sample of Our Clients





Environics Analytics

▲ Founded in 2003; a team of experienced, <u>consultative</u> geodemographers



Jan Kestle, President (former president of Compusearch)



Dr. Tony Lea, Senior Vice President & Chief Methodologist (creator of geodemographic cluster systems for 30 years)



Dr. Doug Norris, Senior Vice President & Chief Demographer (former Director General, Social & Demographic Statistics at Statistics Canada)

...plus 35 others, many with 10+ years of geodemographic experience



Objectives of Project

- Estimate population by age for July 1, 2009 and then estimate possible voters
- Project populations by age for December 31, 2012 and project possible voters
- Do both for the HRM as a whole and also at the smallest spatial scale at Dissemination Area (DA) level
- ▲ Aggregate the DA results up the 23 HRM "2008 Electoral Districts"
- ▲ Explain what you did



Data Inputs

From Environics Analytics

- ▲ At DA and larger geographies for all of Canada we have all of the census data from the (last) 2006 census, including households, population and persons over 18
- ▲ From our Demographic Estimates and Projections (DEP) we have our own product of DA level estimates and projections of households, population by sex and 5 year cohorts and for every year up to 2020
- A large amount of supplementary data and information about demographic, economic and population trends in Canada and regionally
- ▲ TeleAtlas street and boundary files, land use and transportation infrastructure datasets were also used
- ▲ We obtained vacancy rate data from CMHC for regions that include several Census Tracts
- ▲ A wide range of 'boundary files' for mapping and analysis



Data Inputs

From HRM

- ▲ GIS files showing roads, streets and highways
- Civic address points very comprehensive, except does not have reliable upto-date number of dwelling units for each point
- Building permits database for after 2006 by address (shows the number of units at address involved before and after alterations)
- "Subdivision database" planned or expected new developments relating to 2009 - 2012 for type of development and (often) the number of expected units by location (address)
- ▲ A wide range of detailed databases and mapped data on HRM land uses and land use restrictions to help us assess where growth may take place in the period 2009 to December 31, 2012: water bodies, preservation areas of all kinds, zoning areas, park lands of various kinds list on next page
- ▲ Boundary files for DAs etc in own map projection
- ▲ Detailed map of the 23 HRM "2008 Electoral Districts" in current use now
- Wisdom/advice and support especially on location knowledge and data from a range of very informed persons on the HRM team



Data for Project from HRM

Transportation and Utilities

- Future transportation routes
- Current road network
- Transmission lines
- CN-Rail
- Service requirements

Vegetation and Natural Features

- Wetlands, swamps
- Lakes
- Streams
- Vegetation cover

Administrative Boundaries

- 2006 Dissemination areas
- 2006 Dissemination blocks
- 2008 Electoral districts

General Land Use

- Cemeteries
- Provincial parks
- Municipal parks
- Schools
- Golf courses
- Universities and colleges
- Land use points

Detailed Land Use Data

- Civic address points
- Conceptual growth centres
- Large infill potential areas
- Master plan areas
- Vacant parcels
- Greenfield areas
- Business parks
- General future land use areas



Overview of Methodology

PART A: Large Geographical Scale Analysis

1. Study the growth of **HRM as a whole** in the past and using these data and also data on trends estimate July 1, 2009 population and 2012 end of year population for HRM as a whole

PART B: Small Scale Analysis

- 2. Estimate of July 1, 2009 dwellings/households, population by age and possible voters **at the CT and DA levels of geography**
- 3. Projections of December 2012 dwellings/households, population by age and possible voters at the CT and DA levels of geography



Methodology (A) – Large Scale

PART A: Large Scale Estimates

The components of growth are:

▲ Birth rates, death rates

and

- ▲ Interprovincial In-migration
- ▲ Interprovincial Out-migration
- Immigration
- ▲ Emigration
- ▲ Net Temporary Emigrants
- Returning Emigrants
- ▲ Net Change in the Number of Non Permanent Residents
- ▲ The estimates from our municipality level estimates were used as "controls" in our small geographical scale analysis to follow



. . .

Net interprovincial migration

Net international migration

=> Rate of natural increase

Methodology (A) - Large Scale

- ▲ Step 1- Estimate of Halifax population in 2009 adjusted for Census undercount
 - Preliminary estimate of population for all of Nova Scotia was combined with past trends for Halifax and rest of Nova Scotia to obtain a 2009 estimate of population for Halifax
 - Age distributions based on good STC estimates for the CD of Halifax for 2008
- ▲ Estimate of Halifax Population in December 2012
 - The net result was population growth of an average of 1% a year comparable to estimated growth over the period 2007-2009 and higher than growth over the period 2001-2007



The Relevant Geographic Regions and Areas

- ▲ The key regions and areas are:
 - HRM
 - Electoral Districts
 - Census Tracts (CTs) (nest inside regional municipalities)
 - Dissemination Areas (DAs) (nest inside CTs)
 - Dissemination Blocks (DBs) (nest inside DAs)
- ▲ Next slides here show maps of the HRM boundaries:
 - 23 "2008 Electoral Districts"
 - 88 CTs
 - 576 DAs
 - 3166 Dissemination Blocks, DBs



Map of Relevant Small Areas



LYTICS

Α

Α Ν

Zoom-In from Previous Map





Methodology (B) – Small Scale

Estimates For July 1, 2009

- 1. Estimate 2009 dwellings/households at the CT and DA levels of geography
- 2. Estimate 2009 population by age at the DA level of geography
- 3. Estimate 2009 possible voters at the DA level of geography
- ▲ **Dwellings/households** are the demographers natural first step
- ▲ The numbers we estimated for dwellings in 2009 were about 85% driven by your own municipal building permits data (so these are necessarily more reliable than the 2012 projections simply because of the reliable data that we used)
- ▲ The other data driving growth from 2006 census were EAG's regular growth assessment variables which include special variables we construct on:
 - growth rates by each small area (CT and then DA) from the past
 - nearness to past growing areas and to dense areas
 - nearness to transportation infrastructure
 - influenced by current density relative to the density norm in similar areas
 - nearness to quantities of vacant land or low population density land (as long as it was known to be zoned for development)
- Elaborate QC process that involved looking at areal photography (including Google Earth) to see if it looked like the land uses would be able to permit the growth that had been algorithmically assigned to this area
- ▲ This areal photography QC process was also used for 2012 projections



ASIDE: Creation of CT/DA Growth Attraction Score

- ▲ The synthetic variables created by EAG are of three types: density, mathematical potential, and gradient.
- ▲ The density measurements created are based on households and population per possible usable lands (i.e. only land area where residential units can be legally constructed is considered usable).
- ▲ Mathematical potentials are computed using different decay functions to estimate proximity to activity centers. Thus, these mathematical potentials are computed for a variety of variables: population, households, population and household growth, transportation infrastructure, and employment/business.
- ▲ The last synthetic variable, gradients are analogous to the slope of terrain and are computed by measuring a DA or CT's growth rate against its neighbours.
- ▲ All of these synthetic variables along with other non-synthetic variables are used to generate an 'attraction score' for growth for each CT or DA. Once the attraction score is generated the estimation and projection process can begin.
- These synthetic variables were created for both mid-year 2009 and end of year 2012



2009 Population, Ages and Voters

- ▲ Given 2009 dwellings = households(DA level here only) our first estimate of pop came from multiplying household size by the number of households
- ▲ The size multiplier came from 2006 census "pop per household ratio but reduced a small amount because of overall trend; if new households were expected to be younger ones then the household size was increased a small amount
- ▲ Ages of population; we had CD level Statistics Canada estimates for age distribution for 2008 and our own age distribution data at the CSD level
- ▲ We projected 5 year age cohorts based on trends from these from 2006 (adjusted) to 2009; then adjusted all these counts so they added up to our earlier estimated age distribution for HRM
- ▲ To get possible voters (population 18+) we needed to remove non-Canadians 18+ from the general population estimates
 - immigrants have a different age distribution and we had HRM data re this
 - we had data on ages of Canadian citizens so we were able to do a good job estimating ages by population types ... to estimate possible voters



Possible Voters Map 2009 – Zoom-In





Methodology (B) - 2012

For December 31, 2012 ... we first did households

- 1. Project 2012 dwellings/households at the CT and DA levels of geography
- 2. Project 2012 population by age at the DA level of geography (then ages)
- 3. Project 2012 possible voters at the DA level of geography
- ▲ The reference year was the one we just completed 2009
- ▲ The basis for the allocation of the pre-projected new HRM growth was based on these variables:
 - 1. The planned new developments in the municipal database
 - 2. All of the data provided by HRM that related to this (including where growth was taking place provided by the HRM staff
 - 3. All of the data used for 2009 allocation that included transportation infrastructure, previous rates of growth, previous densities relative to overall densities in HRM
- ▲ We used probabilities to assign the new 2009 to 2012 HRM growth from subdivisions database (share of the households to be assigned) to CTs and later DAs :
 - 100% of the dwellings to those areas that were deemed very likely to have been built before the end of the period
 - 50% to those areas that were deemed to have some probability of having been built, and,
 - 25% to those that were deemed unlikely to have been fully built before the July 1, 2009



Methodology - 2012 Populations, Ages and Voters

- Projected populations by multiplying household by area specific household size - as done for 2009 but for 2012 we used a projection
- Then all adjusted to add up to the HRM control numbers for total for 2012
- ▲ Ages of population; we had age cohorts projections from EAG Demographic Estimates and Projections (DEP) data for 2012 to use as a guide
- ▲ To project voters (population 18+) we needed to remove non-Canadians 18+ from the general population estimates
 - immigrants have different age distribution and had HRM data re this
 - we had our population ages projections so we were able to do a good job of projecting voters here
- ▲ At this point the results available for households, population, (pop by age) and possible voters at the DA level



Moving the DA Level Estimates to Districts

- ▲ When we were finished the estimates for 2009 and the projections for 2012 at the DA level, it was necessary to re-state the results for the 23 current 2008 Electoral Districts
- ▲ The challenge is that the Districts are not made up of DAs (wholly in or out) so that DAs would have to be split in order to accomplish the task
- So we made use of the little Stats Canada areas that make up each DA – called Dissemination Blocks
 - in fact typically road bounded physical blocks
- ▲ The error caused by slitting blocks is less than the error caused by splitting DA's (404 Blocks split of 3163 12.8%)
 - but only 148 (1%) or 73 (10%) split in a significant manner
- ▲ No ages and other census data for blocks so we reasonably assumed the other data we needed was distributed evenly within Blocks



Districts, DA's and Blocks





The Zones of Possible Error in Aggregation





Split Blocks





Results For the 23 2008 Electoral Districts

- ▲ The results that will be shown next go in the order that we derived them households, populations, (by ages), and finally possible voter
- ▲ Recall that the two times involved were July 1, 2009 and December 31, 2012
- ▲ There are 3 sheets of paper in the 11 by 17 inch handout package:
- ▲ The first page shows the (households, population, possible voters) counts for these:
 - 2006 census counts
 - 2006 Estimates (from EAG's standard product)
 - Our 2009 Estimates
 - Our 2012 Projections
- ▲ The second page shows all relevant differences
 - 1. 2006 estimates 2006 census
 - 2. 2009 estimates 2006 estimate

- d
- 4. 2012 projection 2006 estimate
- 3. 2012 projection -- 2006 census
- 5. 2012 projection -- estimate 2009
- ▲ The third page shows the "percentage differences" for the same set of 5 breaks shown above for differences
- ▲ The "2006 census possible voters" are closely approximated by citizens 18 plus



Results For the 23 2008 Electoral Districts

Please look at the 11 by 17 inch package of three pages as I take you through it



Questions



Households Estimates 2009 - Whole HRM





Households Estimates 2009 – Zoom-In





2009 Population Estimates





2009 Population Estimates – Zoom In





2009 Possible Voters Estimates





Dissemination Block Allocation - Splits



