# Comparison between Gravity and Destination Choice Models for Trip Distribution in Maryland

- 3 4 By 5 6 7 Sabyasachee Mishra, Ph.D., P.E. Assistant Professor, Department of Civil Engineering, University of Memphis, Memphis, TN38152, 8 Phone: (901) 678-2746, Email: smishra@memphis.edu 9 10 Yanli Wang (Corresponding Author) 11 Ph.D Assistant, Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji 12 University, Shanghai, China 201804, Phone: +86-135-0178-0903, email: vanliwang2012@gmail.com 13 14 15 Xiaoyu Zhu, Ph.D. Faculty Research Associate, National Center for Smart Growth Research and Education, University of Maryland, College Park, MD 20742, Phone: +1-301-405-8000, e-mail: 16 xyzhu@umd.edu 17 18 Rolf Moeckel, Dr.-Ing. Supervising Research Engineer, Parsons Brinckerhoff, 6100 NE Uptown Boulevard, Suite 700, 19 Albuquerque, New Mexico 87110, Phone:+1-505- 878-6553, email: moeckel@pbworld.com 20 21 22 Subrat Mahaparta 23 Office of Planning and Preliminary Engineering, Travel Forecasting and Analysis Division, Maryland State Highway Administration, 707 North Calvert Street Baltimore, Maryland 21202, email: 24 25 smahapatra@sha.state.md.us 26 27 28 Word Count:5,698 29 Number of Tables:3 Number of Figures:4 30 Total Count:  $= 5,698 + (7 \times 250) = 7,448$ 31 32 Date Submitted: November 15, 2012 33 34 35 36 Submitted for Peer Review and for Compendium of Papers CD-ROM at the Annual Meeting of the
- 37 Transportation Research Board (TRB) in January 2013

# 1 ABSTRACT

2 Gravity models have been widely used in the trip distribution stage of travel demand models over decades.

Recent advancements in research suggest that the destination choice model is a preferred trip distribution
 method rather than the gravity model. In this paper, a comparison of destination choice model and gravity

model is presented using a real case study applied within the Maryland Statewide Transportation Model. The gravity model is formulated as a function of travel time and origin-destination's production-attraction ability and the destination choice model is based upon the utility maximization principle. Household Travel Survey data collected in the Baltimore-Washington region is used to estimate and calibrate the parameters of both models. Considering the wide spectrum of diversity in growth pattern in the

10 Baltimore-Washington survey data, the study area is divided into eight regions to reflect region specific

travel behavior. From the result, it is found that a destination choice model better replicates the observed trip length and origin-destination matrix for home based work (HBW) trips. In contrast, the gravity model

does not succeed in accurately estimating trip attractions when compared to the survey data. The result

shows that a destination choice model performs better than a gravity model in this statewide model. This

research can be used as a tool to assist in choosing a trip distribution method for travel demand modeling.

16 Keywords: trip distribution model, gravity model, destination choice model, utility maximization

# 1 1. INTRODUCTION

2 Modeling trip distribution is critical to traffic demand forecast and transportation planning. Trip 3 distribution is the second component (after trip generation) in the traditional four-step transportation planning model. This step matches trip makers' origins and destinations to develop a matrix that displays 4 the number of trips between each origin and destination (OD). Trip distribution models are intended to 5 produce the best possible predictions of travelers' destination choices on the basis of generation and 6 attraction abilities of each transportation analysis zone and the level of impedance (e.g., generalized cost 7 8 of) between each OD pair (1). Actually, the interactions between origins and destinations are complex and the trip distribution model is a major weakness in the four step model (2). 9

Several trip distribution models have been developed (3–5). The first one was the Fratar Growth model followed by a number of other models including Simple Growth factor model, Furness model, and Detroit model are developed at the same time period (6). The next model developed was the gravity model. Since Casey (7) proposed the gravity model for trip distribution, the gravity model was widely accepted and experienced several improvements such as doubly constrained gravity models (8–11). Till now, it is still the most popular distribution model (4).

However, due to the more diverse personal travels and the more accurate requirements for
forecast, the gravity model shows the limitations as an aggregate model. With the development of logit
and other discrete choice techniques, disaggregate approaches are introduced to travel demand forecast.
Random utility (RU) theory offers the one of the most flexible and powerful paradigm for travel demand
modeling and especially for trip distribution (12). RU destination choice model is proposed in the 1980s
(13–17). Many researchers study its application and integration with travel demand models (18–22).
Since this model requires a large amount of data, it still not widely applied in large-scale models.

Because of the improvements in technology and data collection, accuracy of the model is more 23 24 important. Currently, the expanded data available have made various models more comparable in terms of 25 forecasting capability. The objective of this paper is to provide a comparison of a doubly-constrained gravity model and a multinomial logit destination choice model integrated in a large-scale model. 26 27 Maryland (Washington DC and Baltimore region) is selected as the study area, and the survey data are collected to calibrate parameters and compare with the fitting results of both gravity model and 28 destination model. This paper provides the comparison results in a statewide application and targets to 29 claim the better model. It will effectively promote the application of the destination model in large-scale 30 31 trip distribution models.

The remainder of this paper is organized as follows. In the next section, a literature review shows the applications of gravity model and destination choice model and their limitations. Section 3 is compromised of the formulation of the double constrained gravity model, and the destination choice model. Section 4 describes the study area and data sets. Section 5 presents the calibrated parameters, the fitting results, and the comparison result to the real data for both methods. Finally, Section 6 highlights the main conclusions and recommendations achieved from this study.

# **2. LITERATURE REVIEW**

39 The literature review is presented in three themes. The first two themes discuss about the arguments from 40 researchers on gravity model and destination choice model. The third theme discusses the disagreement

41 that occurred in the distribution models.

# 1 2.1 The Development and Application of Gravity Model

The gravity model was initially proposed in 1950s for trip distribution (7), and some researchers make some important improvements in 1960s (23). Afterwards, many researchers examined and calibrate the gravity model (10), (24–26). Until now, there still some studies working on the parameters or the improvements for the factors(1), (4), (11), (26). The gravity model is the most popular method and is extensively adopted in most MPO and statewide travel demand models in US (4), (5), (27).Gravity model assumes that the trips produced at an origin and attracted to a destination are directly proportional to the total trip productions at the origin and the total attractions at the destination.

9 In the calibration of the gravity model, the model generally fits well when the calibrated 10 frequency distribution is statistically close to the observed frequency distribution of the base year trips(1-11 3), (25). The comparison of trip length frequency distributions does not necessarily provide enough 12 information for assessing how well the trip interchanges match(24).Therefore, a "good fit" of trip length 13 frequency distributions does not guarantee accurate predictions of future zone-to-zone trip interchanges. 14 Because the predicted trips will be used in the following phase of the demand model, it is important to 15 produce reliable and consistent origin-destination (O-D) trip interchanges.

16

# 17 **2.2** The Development and Application of Destination Choice Model

The destination choice models based on random utility theory and discrete choice modeling theory are proposed and applied to trip distribution in the 1980s (13–17). This approach offers the most flexible and powerful paradigm for trip distribution modeling. The model incorporating the travel time and socioeconomic variables in the utility or impedance function is different from the gravity model in concept which uses travel time as impedance to determine the probability of trip making.

Since destination choice model was proposed, a lot of studies have been done on the improvements and application. Ben-Akiva and Lerman (16) have developed a joint destination choice and mode choice model using a logit formulation for work and non-work trips. Thill (28), (29) developed choice set formation for destination choice model. Bhat (30) formulated a disaggregate attraction-end choice model for the practical application, and the model was integrated to the transportation planning model in Portland, Oregon. The application of destination choice model is still widely studied (20–22), (31–34), and the method is also applied to tourism destination study (20).

30 However, though the destination choice model well addressed the capability to accommodate population heterogeneity, it still not widely used in large-scale. Since destination choice model is 31 32 estimated at disaggregate level, impedance variables interact with individual's demographic and socioeconomic characteristics. Therefore a destination choice model with continuous attributes interacting with 33 impedance cannot be applied into practice since individuals' continuous attributes are unavailable at TAZ 34 (Traffic Analysis Zone) level unless an activity-based model is applied to simulate each individual's 35 travel behaviors. And also the spatial aggregation error will occur in both model estimation and 36 37 application when destination choice model is developed at individual level but applied at TAZ level (35).

38

# 39 **2.3 Researcher's Debate**

The debate between gravity model and destination choice model started since the destination choice
model appeared. There are some other researchers working on this topic before (3), (35–38).

42 The advantage of destination choice models over gravity models is the presence of the 43 agglomeration effects and hierarchy in destination choices (15), (13) or incorporating agglomeration effects related to trip chaining and spatial autocorrelation and competition (39). Chow et al. (40)
mentioned the other advantages of destination choice model (such as simultaneously modeling the effects
of land use on travel behavior) and a disadvantage of gravity model (unable to evaluate many policy
issues for long-range planning).

5 While some researchers who worked on the gravity model confirmed that the gravity model can incorporate or stratified by the socioeconomic factors, and adjust the friction factor which is not 6 7 necessarily linear. Therefore, market segments can also be differentiated in gravity model if friction 8 factors are calibrated to match trip length distribution of specific market segment (population). In addition, 9 due to the suspicion of spatial aggregation errors, most practitioners still use gravity model (35). Also 10 there are only a few researchers compared the fitness of the two distribution models using real data (38) and even fewer people compare the different results of the two models using statewide data and model. 11 The destination model is inconclusive to be better than gravity model. 12

In this paper, the Maryland (Washington DC and Baltimore region) data is collected, and the statewide model is developed to get the results of destination choice model and gravity model so as to compare and evaluate their performances. The objective is to help practitioners to achieve a better understanding of these two alternative modeling methods and make their own decision for trip distribution model

18

# 19 **3. METHODLOGY**

In this section, the formulations of gravity model and destination choice model are presented. The gravity
 model aims at friction factors between each OD pair and for all trip purposes at an aggregate level.
 Destination choice model predicts the probability of choosing any destination zone discretely from
 maximum utility point of view.

24

#### 25 **3.1 Gravity Model**

Here, the gravity model we applied is stratified by trip purpose and Home-based trip purposes stratifiedby income category (five income groups). The basic gravity model formulation is shown below:

$$T_{ij}^{k} = P_i^{k} \cdot A_j^{k} \cdot F_{ij}^{k} / \sum_{j} (A_j^{k} \cdot F_{ij}^{k})$$
<sup>(1)</sup>

28 Where:

29  $T_{ij}^{k}$  = trips for purpose 'k' between production SMZ 'i' and attraction SMZ 'j'

30  $P_i^k$  = productions for trip purpose 'k' in Statewide Modeling Zone (SMZ)'i'

31  $A_j^k$  = attractions for trip purpose 'k' in SMZ 'j'

32  $F_{ii}^{k}$  = friction factor for trip purpose 'k' between SMZ 'i' and 'j'

Friction factor  $F_{ij}^k$  is a function of composite impedance incorporating auto and transit travel time and cost components (e.g., mode choice logsums). The function of friction factors takes the following form: Mishra et al.

$$F_{ii}^{k} = \alpha \cdot (CT_{ii}^{k})^{\beta} \cdot exp \left(\gamma \cdot CT_{ii}^{k}\right)$$
<sup>(2)</sup>

1	Where:	
2	$CT_{ij}^k$ = Composite Time for purpose 'k' between SMZ 'i' and 'j'	
3	A =10,000,000	
4	$\beta$ and $\gamma$ = calibrated coefficients	
5		
6	$CT_{ij}^k$ , the composite time for purpose 'k' between SMZ can be defined as	
7		
	$CT_{ij}^{k} = (1/(\frac{1}{HT} + \frac{x}{TT}) + y \cdot TL/vot) \cdot Adj$	(3)
8	Where	
9	HT = highway time, minutes (including O+D terminal time)	
10	TT = total transit time, minutes (walk+wait, best walk-access path)	
11	TL = highway toll, cents	
12	<i>vot</i> = value of time, cents/minute	
13	x, y = coefficients (vary by income and purpose)	
14	Adj = Calibrated coefficients	
15	The weakness of the gravity model is the inability to take account of other impe	dance

16 components rather than travel time and cost, such as employment, and household information.

17

#### 18 **3.2 Destination Choice Model**

The destination choice model predicts the probability of choosing any given zone as the trip attraction end. The model was estimated in a multinomial logit form using the ALOGIT software. These models are preceded by the trip production models, which forecast the number of productions by zone for different trip markets, chiefly identified by purpose and household income level. The destination choice models include mode choice logsums, distance terms, zonal employment, household characteristics and region geographic characteristics.

The household, trip, zone related variables were examined and proved to be significant on many different purposes. By allowing for the inclusion of multi-modal accessibilities and several other region and trip market terms, the destination choice framework helps explain variation in travel across the state that was difficult to explain with a single gravity model impedance function.

The utility  $(U_{ijn})$  of choosing a trip attraction destination (j) for a trip (n) produced in zone (i) is given by:

$$U_{ijn} = S_j + \partial \times L_{ij} + \sum \beta^k \cdot D_{ij}^k + \sum \beta^k \cdot D_{ij}^k \cdot N_n^k + \sum \beta^k \cdot Z_j^k + C_{jn}$$

$$\tag{4}$$

31 Where,  $S_j$  is the size variable for destination zone j,  $L_{ij}$  is the mode choice logsum between zone 32 pair ij,  $D_{ij}^k$  represents the various distance terms (linear, log, squared, cubed and square root),  $N_n^k$  represent person, household or production zone characteristics for trip n and is used for creating interaction variables with distance terms,  $Z_j^k$  represents attraction zone characteristics (other than the size term), and  $C_{jn}$  is a correction term to compensate for the sampling error in the model estimation (i.e., it represents the difference between the sampling probability and final estimated probability for each alternative).

5 The size variable may consist of several different terms; employment categories in addition to 6 households. Weights ( $\beta^k$ ) for each term in the size variable were estimated along with all other model 7 parameters as follows, where  $E_j^k$  is employment of type k in zone j:

$$S_j = \log\left(\sum \beta^k \cdot E_j^k\right) \tag{5}$$

8 Since the scale of the size term is arbitrary, one of the  $\beta^k$  coefficients is always set to 1.0. 9 Analternative and equivalent specification of the size variable, implemented in ALOGIT is

$$S_{j} = \log\left(\sum \exp\left(\lambda^{k}\right) \cdot E_{j}^{k}\right) \tag{6}$$

10 ALOGIT reports the value of  $\lambda^k$ , instead of reporting directly the value of  $\beta^k$ . For this reason, the 11 estimated size term coefficients may be negative; the actual coefficients are of course always positive, 12 consistent with theory.

A combination of distance terms is used in the utility such that the composite distance utility function is monotonically decreasing. These distance terms are used to closely approximate the shape of the trip length frequency distribution. The distance-related disutility may be capped at a chosen maximum value, to maintain a reasonable probability of selecting far away destinations. The distance cap was established during model estimation at 30 miles, and may be adjusted during model calibration to ensure that the model reproduces the tail of the trip length frequency distributions. In addition, the choice probability needs to be adjusted with sampling correction factors  $ln\left(\frac{n_i}{N \times a(i)}\right)$  by the following formula:

$$P(i) = \frac{exp\left[V_i + ln\left(\frac{n_i}{N \times q(i)}\right)\right]}{\sum_{j \in D} exp\left[V_j + ln\left(\frac{n_j}{N \times q(j)}\right)\right]} = \frac{\left(\frac{n_i}{N \times q(i)}\right) \times exp\left(V_i\right)}{\sum_{j \in D}\left(\frac{n_j}{N \times q(j)}\right) \times exp\left(V_j\right)} = \frac{\left(\frac{n_i}{q(i)}\right) \times exp\left(V_i\right)}{\sum_{j \in D}\left(\frac{n_j}{q(j)}\right) \times exp\left(V_j\right)}$$
(7)

21 Where,

22  $i \in C$  = unique alternatives from the full set

23  $j \in D \subset C$  = unique alternatives from the sample

24 q(i) = selection probability (probability to be drawn)

25  $n_i$  = selection frequency in the sample

- 26  $N = \text{sample size} \sum_{i \in D} n_i = N$
- 27  $V_i$  = utility of a choice alternative
- 28 P(i) = choice probability

# 29 4. DATA ANDCASE STUDY AREA

To demonstrate the application of the proposed methodology we used state of Maryland as a case study. 30 31 The state of Maryland consists of 23 counties, one independent city and with a total population of 5.8 32 million and a total employment of 3.4 million in the year 2010 (41), (42). In the state of Maryland two Metropolitan Planning Organizations (MPOs): Baltimore Metropolitan Council and Metropolitan 33 34 Washington Council of Governments (MWCOG) in the past maintain travel demand models. In 2007-2008, the Baltimore Metropolitan Council (BMC) on behalf of the Baltimore Regional Transportation 35 Board, teamed with the Transportation Planning Board at the Metropolitan Washington Council of 36 Governments (MWCOG) to conduct a household travel survey (HTS) in both the Baltimore and 37

- 1 Washington regions.Data for the survey was collected from randomly selected households in the
- 2 Baltimore and Washington DC region. Map of the data processing regions is shown in Figure 1.





#### FIGURE 1 HTS Data Processing Regions

5 Each household completed a travel diary that documented the activities of all household members on an assigned day. Demographic information was also collected. The surveys have been stored in a 6 database, which contains records for approximately 4,500 households, 10,000 persons, 49,000 trips, and 7 8 6,000 vehicles. The HTS data consist of four separate files - a household, person, trip and vehicle file. In 9 the data analysis, it was found that travel pattern was significantly different. Figure 2 shows the trip length for different regions for Home Based Work Purpose. Each region was assigned based on the FIPS 10 code of the home location of the household record corresponding to the trip. The survey data was a 11 sample for households and trips. Appropriate weighting factors were used to estimate the whole 12 13 population.

14 In the doubly constrained gravity model, trip production and attraction data was used to estimate trip distribution between regions. Trip rates are derived from HTS for household by workers and size 15 cross classified by income. Corresponding land use data were derived using cooperative forecast of all the 16 17 state agencies included in the study area. Trip production and attraction were estimated using cross classification method using trip rates and land use data. In destination choice model, the probability of 18 19 choosing any given zone as the trip attraction end is derived using utility maximization principle. The 20 model coefficients along with HTS population data are used to derive trip ends. For comparison purposes in the result section, survey (population) trip ends are compared with similar counterparts in the gravity 21 22 and destination choice model.

# **1 5. RESULTS ANDDISCUSSION**

In this section, we present the results of gravity model and destination choice model and compare the differences between the outputs of these two models in MSTM. Gravity model calculates 18 friction factors including three home-based purposes across five income groups, and three other trip purposes. Destination choice model is applied to five trip purposes (not including home-based school) and coefficients such as distance, income are estimated in the model.

7

#### 8 5.1 Gravity Model Results

9 In the gravity model, there are six types of trip purposes: HBW, HBS, HBO, HBSc, NHBW, and OBO. Home-based purposes (HBW, HBS and HBO) are stratified by five income groups. Friction factors 10 functions incorporate segmented distance terms to facilitate calibration to target trip length frequency 11 12 distributions over the longer distances encompassed by the statewide model. Friction factors parameters 13  $(\beta, \gamma)$  and Composite Time parameters (x, y, VoT, Adj) developed for the gravity model is displayed in Table 1. The  $\beta$  and  $\gamma$  parameters are calibrated for MSTM while the other parameters are consistent with 14 15 the BMC model. When the x and y are zero, the composite time impedance is simply highway time. The combination of greater variance in trip rates by area, greater market segmentation by income and 16 inclusion of segmented distance terms in the trip distribution impedance, should minimize the need for 17 18 trip distribution adjustment factors.

19

**TABLE 1 Friction factors and Composite Time Parameters** 

D	т	0		1		V/T	A 1'
Purpose	Income	β	γ	Х	У	Vol	Adj
	1	0.6	-0.174	0.3	0.9	8.4	1.555
	2	0.7	-0.154	0.02	0.9	25	1.401
HBW	3	0.8	-0.116	0	0.9	41.7	1.261
	4	1.1	-0.108	0	0.9	50.4	1.247
	5	0.8	-0.102	0	0	106.4	1.259
	1	0.1	-0.66	0	2.1	8.4	1.552
	2	0.1	-0.34	0.02	0.7	25	1.142
HBS	3	0.1	-0.3	0.01	0.7	41.7	1.057
	4	0.1	-0.33	0	0.7	50.4	1.038
	5	0.1	-0.36	0	0.7	106.4	1.040
	1	0.1	-0.46	0	0	8.4	1.548
	2	0.1	-0.21	0.05	1.35	25	1.218
HBO	3	0.1	-0.214	0.02	1.35	41.7	1.101
	4	0.1	-0.26	0	1.35	50.4	1.045
	5	0.1	-0.23	0	1.35	106.4	1.086
HBSc	all	0.1	-0.36	0	0	45.2	1.546
NHBW	all	0.1	-0.15	0	1.25	45.2	1.020
OBO	all	0.1	-0.26	0	2.1	45.2	1.054

20

Calibration targets for the gravity model also include average trip length and trip length frequency distribution. The calibration targets are initially developed in the conventional way, using model estimated travel times (skims). It is found that the skim travel times resulted in significantly shorter distances then reported in the surgery. Appendix the skim matrix tands to underestimate conception and

24 distances than reported in the survey. Apparently, the skim matrix tends to underestimate congestion, and

- therefore suggested that travel times are shorter than found in reality. In other words, in the same amount of time, people can travel further in the model than they can in reality.
- 3

#### 4 5.2 Destination Choice Model Results

The calibrated coefficients implemented in the destination choice model are shown in Table 2. The results
include five trip purposes: HBW, HBS, HBO, NHBW and OBO. The variables include trip characteristics
(mode choice logsum, distance polynomial), household characteristics (income, auto ownership), and
location characteristics (like CBD indicator, region-specific indicator, bridge crossing, and the size term).

9 The result for HBW and HBS are discussed in detail below.

#### 10 Home-based Work (HBW) Model Estimation

The first model estimated was the home based work purpose. There are 20,626 HBW trip records in the 11 survey file. The model specification was built incrementally, starting with a utility function that included 12 13 only the mode choice logsum, and adding distance terms, size terms, and other trip attributes one at a time, in order to better match observed trip patterns. Various specifications with a capped distance disutility 14 were explored. The purpose of the cap in estimation is to reduce the influence of very long but infrequent 15 trips on the distance polynomial coefficient estimates. Approximately 90% of the HBW trip observations 16 exhibit a distance shorter than 30 miles (see Figure 2). The final HBW model was estimated with a 25 17 mile distance cap. Estimation runs also tested trip length differences among socio-economic variables and 18 home residential location, and the attractiveness of the two CBD areas. A key focus was on the segmented 19 20 distance term to match the short distance portion of the observed trip length frequency curve.

21 From the result in the first column of Table 2, we observed the important findings in the variables. 22 The mode choice logsum coefficient is 0.58, consistent with theory and with the expectation of relatively elastic demand. The coefficients for the distance polynomial are all significant, and the combined total 23 24 utility with respect to distance decreases monotonically with distance, as expected. Household income was interacted with the linear distance term but is not statistically significant. The positive and significant 25 intrazonal coefficient indicated a preference for destinations in the home zone, all other things equal. 26 27 Following the trip generation segmentation, trips are stratified into five household income groups. The income coefficients are expressed relative to the lowest income (<\$30K), which was given an income 28 29 coefficient of zero. The coefficients on the other income categories were positive and significant, with a 30 steadily increasing coefficient value as the income level increases. This result shows that persons from higher income households are likely to make longer work trips. The effect of auto ownership is highly 31 correlated with household income. No auto ownership effects were kept in the final estimated model. 32

**TABLE 2** Calibrated Coefficients for Destination Choice Models

	Trip Purpose				
Explanatory Variable	HBW	HBS	HBO	NHBW	OBO
Mode choice logsum	0.5769	0.8000	0.8420	0.9078	0.8000
Distance	-0.4383	-0.3986	-0.5788	0.0978	-0.2241
Distance Squared	0.0137	0.0166	0.0261	-0.0032	0.0106
Distance Cubed	-0.0002	-0.0004	-0.0005		-0.0002
Log of Distance	0.7066	-0.9034	-0.4212	-1.5665	-1.0944
Intrazonal indicator variable	1.2038		0.6633	0.7228	0.6311
Income X Distance interactions					
Income (<30K)					
Income (30-60K)	0.0176	0.0162	0.0345		
Income (60-100K)	0.0470	0.0255	0.0357		
Income (100-150K)	0.0606	0.0263	0.0357		
Income (150K+)	0.0697	0.0263	0.0357		
Size Term (exponentiated)					
Other Employment	1.0000		0.3052	0.4271	0.1470
Retail Employment	1.0134	1.0000	0.1878	1.0000	1.0000
Office Employment	0.2904		0.0446	0.4992	
Industrial Employment	0.3585				0.0874
Households			1.0000	0.2825	0.3243
Baltimore CBD (Region 1)	-0.0362				
Washington DC CBD (Region 2)	-0.0882				
Bridge Crossing indicator	-0.3013	-1.2928	-0.8054	-0.5280	-0.9768
Baltimore Semi-Urban (Region 3)	-0.0269				
Wash.DC Semi-Urban (Region 4)	-0.0422				
Baltimore Suburban (Region 5)	-0.0350				
Wash. DC Suburban (Region 6)	-0.0255				
SE Maryland and Halo (Region 7)	-0.0255	-0.0100	-0.0100	-0.0100	-0.0100
SW Maryland and Halo (Region 8)	-0.0350	-0.0100	-0.0100	-0.0100	-0.0100
Distance Cap	25	30	30	30	30
Distance Constants					
0-1mile	0.7729	1.7660	1.4007	0.2417	2.2193
1-2 miles	0.0000	1.9110	0.5347	0.0140	1.1874
2-3miles	-0.1059	1.2765	0.1937	-0.0396	0.6676
3-4 miles	-0.3221	0.8224	0.1937	-0.0396	0.6676
4-5 miles	-0.1424	0.7539	0.1937	-0.0396	0.6676
5-6 miles	-0.1424	0.2023	0.0000	0.0000	0.0000
6-7 miles	-0.1000	0.0721	0.0000	0.0000	0.0000

#### Mishra et al.

1 Location indicators are proved to have significant impacts on destination choice. The size term comprises retail, office, industrial and other services employment. All employment categories exhibited 2 significant coefficients. CBD indicator variables were used to explore the attractiveness of a destination 3 4 in either the Baltimore or Washington DC CBDs. Both CBD variables exhibited negative coefficients that 5 are unintuitive because the CBDs are major attractors. However the attractiveness of the CBDs may already be captured in the mode choice logsums or size terms. The CBD variables were dropped from the 6 7 final model. The effect that the Potomac River has on discouraging trips across was examined with a 8 bridge crossing variable. As expected, its coefficient is negative and significant, indicating that bridge crossings are not desirable. For HBW trips, a bridge crossing is equivalent to 12 additional minutes of 9 travel time. 10



11

12

FIGURE2: HBW Observed Trip Length Frequency Variation by Region

13

# 14 Home Based Shop (HBS) and Home Based Other (HBO) Model Estimation

The expanded sample size for home based shop trips is 3,812,775 observations, and the expanded sample size for home based other trips is 7,684,294. The best model estimated for HBW trips was used as the starting point for HBS and HBO, without the inclusion of the regions interacted with distance. The disutility of distance was capped at 30 miles. Following results are concluded from these two models.

19 The mode choice logsum coefficient for HBS was consistently estimated at a value greater than 20 1.0, which is outside the theoretically acceptable range, and therefore was constrained to a value of 0.8. 21 The mode choice logsum coefficient of HBO was estimated at a value of 0.8, which is a reasonable result. For both HBS and HBO, the distance, distance cubed, and log of distance coefficients were all negative 22 23 and significant. And the distance squared term was positive and significant. The total disutility of distance 24 decreases monotonically with distance, as expected. The intrazonal coefficient for HBS was negative and 25 became insignificant when the logsum coefficient was constrained. The intrazonal coefficient for HBO was positive and significant. 26

The household income coefficients for HBS were positive and significant. The two highest income categories were combined into one to obtain a monotonic progression. The household income coefficients for HBO were positive and significant, but did not steadily increase with higher incomes over the five income groups. Therefore, income was collapsed into three categories: less than \$30K, \$30-60K and \$60K or higher.

Retail was the only significant employment category used for the HBS size term. The HBO size
term consists of number of households, retail employment, office employment, and other employment.
The CBD indicator variables for Washington DC and Baltimore were negative in both HBS and HBO
models. Thus, as was the case for HBW, these variables were excluded from the final model. The bridge
crossing coefficients were negative and significant in both models.

11

# 12 Non-Home Based Work (NHBW) and Non-Home Based Other (OBO) Model Estimation

The expanded sample size for the non-home based work purpose was 2,495,400 observations and the expanded sample size for the non-home based other purpose is 4,054,669 observations. The disutility of distance was capped at 30 miles. Results concluded from these two models are discussed below.

The mode choice logsum coefficient estimated for NHBW is approximately 0.9, which is a 16 reasonable result. The mode choice logsum coefficient estimate for OBO was greater than 1; therefore it 17 was constrained to 0.8. In both models, all exhibited significant coefficients, and a combined distance 18 19 decay function that decreases with distance. Also, the intrazonal coefficients were both positive and 20 significant. The income coefficients were not used in NHBW and OBO models because they are not stratified by income. The size term consists of retail employment, office employment, other employment 21 22 and households for NHBW and consists of retail employment, industry employment, other employment, and households for OBO. Similarly to other models, the CBD indicator coefficients for Washington DC 23 24 and Baltimore were negative and excluded from the final model and the bridge coefficient was negative 25 and significant.

26

# 27 5.3 Comparative Analysis

The comparison results between the two models and the survey are discussed, including average trip length, and OD trips. We display the results for HBW OD trips in this discussion because it is representative of the comparison result.

# 31 Average Trip length

Figure 3 compares observed average trip length (ATL) with simulated trip length for each trip purpose. Trip lengths are slightly overestimated, but overall the patterns of the survey are reflected in both gravity

model, and destination choice model. The ATL for survey is the lowest, while destination choice model

are more close to the survey result than gravity model. The ATL comparison demonstrates that the

destination choice model match better than gravity model in the trip distribution stage. While ATL is

widely used to examine the accuracy of trip distribution, this aggregate measurement is not enough. Then

38 we calculate the OD trips between each zone for a more detail comparison.



2 Note: DC Model is destination choice model.

#### 3

#### FIGURE3 Comparison of Average Trip Lengths between Models

#### 4

5 An important reason for not matching precisely the trip length frequency distribution is the fact that the 6 model has to reach the average trip length across the entire model study area, which is much larger than 7 the survey area (FIGURE 1). The diagram in FIGURE 3, however, compares trip lengths in the survey 8 area only. The destination choice model is able to better replicate the average trip length in the survey 9 area while matching the average trip length across the entire MSTM study area. In contrast, once the 10 gravity model was calibrated to the entire model area, the trip length in the survey area is overestimated in 11 every trip purpose.

12

# 13 **OD Trips**

14 In order to compare the fitness of the two models, we provide survey data, and the difference of OD trips

between the two models and the survey. Table 3 shows the OD Trip results for HBW. In Table 3, there are

16 OD trips between eight by eight OD pairs. The last column "Row Total" is comparing the trip origin

17 distribution among the eight regions. And the last row of "Total" is comparing the trip destination

Mishra et al.

distribution. The first row in each origin includes the percentage of trips between regions derived from
survey data. The percentage value in the cell is the distribution of the attraction for the corresponding
origin region. For example, 2.2% (row "1-Survey" and column "2") is the percentage of the survey trips
generated from region1 and attracted to region 2.

#### 5

6

Destination Row Origin 2 3 7 8 **Totals** 1 4 5 6 59.1% 2.2% 29.9% 3.3% 5.1% 0.2% 0.0% 0.0% 8.5% Survey DC\* 67.4% 2.0% 22.4% 2.1% 5.6% 0.0% 6.0% 1 0.3% 0.0% 1.8% 19.4% 1.5% 4.6% 0.3% 0.0% 0.0% Gravity\*\* 72.1% 6.0% 0.5% 3.7% 0.8% 58.1% 3.3% 32.2% 0.4% 1.1% 23.4% Survey DC\* 0.8% 58.7% 2.9% 32.5% 0.5% 3.7% 0.4% 2 1.0% 11.0% Gravity\*\* 0.6% 51.1% 2.6% 40.3% 0.6% 4.6% 0.3% 0.9% 11.0% 5.2% 9.4% 17.3% 60.5% 6.8% 0.3% 0.5% 0.1% Survey 16.1% 3 DC\* 18.2% 3.4% 63.5% 7.1% 7.8% 0.3% 0.5% 0.1% 20.6% 18.2% 3.8% 8.5% 0.4% 0.4% 0.1% 20.9% Gravity\*\* 2.6% 66.6% Survey 0.9% 26.0% 4.3% 59.5% 1.7% 6.3% 0.4% 0.9% 31.4% DC\* 1.9% 0.8% 27.3% 3.9% 0.4% 0.8% 4 60.7% 6.0% 36.9% Gravity\*\* 0.6% 24.7% 3.0% 63.1% 1.9% 7.1% 0.2% 0.5% 36.9% 6.4% 1.8% 14.4% 7.2% 68.9% 0.4% 0.0% 0.8% 8.3% Survey 6.1% DC\* 7.9% 1.6% 12.2% 71.0% 0.4% 0.0% 5 0.7% 7.8% 0.9% 12.2% 4.3% 0.0% Gravity\*\* 8.3% 73.7% 0.4% 0.6% 8.3% 0.2% 0.8% 29.6% 0.4% 48.9% 1.3% 13.4% 5.4% 7.4% Survey DC\* 30.5% 6 0.2% 12.7% 0.8% 0.3% 50.9% 1.0% 4.3% 10.0% 0.1% 0.4% Gravity\*\* 14.5% 0.6% 31.1% 51.8% 0.8% 2.4% 10.0% Survey 0.2% 7.4% 5.4% 6.8% 0.0% 5.7% 74.2% 0.2% 1.6% 7 DC\* 0.2% 8.0% 5.7% 5.8% 0.0% 5.6% 73.5% 0.2% 2.5% Gravity\*\* 0.3% 8.1% 4.9% 4.8% 0.0% 4.3% 87.6% 0.2% 8.9% 67.1% 3.5% Survey 0.0% 6.8% 0.6% 10.3% 1.8% 13.3% 0.1% DC\* 8 0.3% 4.7% 0.8% 9.6% 1.5% 7.6% 0.1% 75.4% 7.4% Gravity\*\* 0.0% 1.0% 0.0% 3.4% 5.8% 0.1% 85.6% 4.0% 10.8% Survey 8.8% 24.3% 15.7% 31.2% 8.0% 7.1% 1.5% 3.4% 100.0% Total DC\* 9.6% 1.2% 23.1% 16.6% 32.0% 5.2% 6.5% 5.9% 2.2%+ 7.2% 18.6% 17.5% 6.4% Gravity\*\* 29.1% 6.0% 6.4% 8.8% 12.8%

<b>TABLE 3</b> Comparison	<b>Model Results with</b>	Survey for HBWOD	Trips
1			

7

Note: DC\*-Destination Choice Model; Gravity\*\*: Gravity Model

Similar to the survey data, percentage of trips from each zone to the other zone is also presented for destination choice and gravity Model in the second row and third row respectively. For example, from zone 1 to zone 2, there are 2.0% trips occurred from destination choice model; and 2.2% from gravity model. When compared to the survey trips, destination choice model appear to have relatively matched trips than gravity model. Similar comparison can be done from each origin region to each destination fegion. It appears for some regions destination choice model overestimates and underestimates in other with no clear pattern.

8 The "Row Totals" from the last column represents percentage of trip produced from region 1. For 9 example the survey data shows 8.5% trips when both destination choice and gravity model show 6%. 10 Similar observation of equal trips from destination choice model and gravity model are found in region 2, 3, and 6. Underestimation and overestimation is also a crucial to observe in Table 3. From the last column, 11 12 we can see the origin distribution is different from survey in zone 2, 7, and 8, with an underestimate in 13 zone 2 and over estimate in zone 7 and 8. Destination choice model provide a closer distribution 14 comparing with survey. And the last row of "Total-DC" or "Total-Gravity" is the difference of total trips 15 compared to the survey data. From the last row, we can find that the production results from destination choice model and gravity models are similar, except zone 7. The attraction distribution in destination 16 17 choice model also matches the survey result better. In general, the results from destination choice model and gravity models are similar (in origin 1, 2, and 6) and close to the survey and destination choice model 18 19 creates a better distribution in origins 3, 4, and 8 than gravity model. Similar to HBW, other trip purposes destination choice model demonstrate better performance in terms of regional OD flows. For brevity all 20 21 other purposes are not shown in the paper.

For home based work trip purpose, flows between regions for both models and comparison with the survey data are presented in Figure 4. Figure 4(a) and 4(b) show the performance of the gravity model and the destination choice model in comparison to the survey data. The X and Y axes show the eight regions, and the Z-axis shows the deviation from the survey, i.e. proportion of model to survey. Values greater than 1 (shown in yellow and red bars) on the Z-axis show region pairs where the model generated more trips than the survey, values below 1 (blue bars) show region pairs that underestimate volumes.

Figure 4(a) shows how some regional flows are not performing well with the gravity model. In comparison, the graphic in Figure 4(b) shows that the destination choice was able to produce a lot more OD pairs that match the survey data, showing particular improvements in these isolates locations which may represent the CBD, or other key trip end attribute. The r-square for the gravity model is 0.47 and the corresponding value for the destination choice model is 0.79. For HBW trip purpose the result shows that the destination choice model performs much better than the gravity model. Discussion on other trip purposes is not shown in the paper for brevity, but a similar trend was found.



Figure 4(a) Regional Flows and Comparison with HTS for Doubly Constrained Gravity Model (r-square: 0.47)





#### **FIGURE 4 Regional Flows and Comparison with HTS**

2

# **3 6. CONCLUSION**

Modeling trip distribution is critical to travel demand forecasting and transportation planning. Trip distribution models are intended to produce the best possible predictions of travelers' destination choices on the basis of generation and attraction abilities of each travel zone and the level of impedance between each OD pair. Gravity model has been widely used in the trip distribution stage of travel demand models over decades. Recent advancements in research suggest that the destination choice model is a preferred trip distribution method over the gravity model. Using Maryland Statewide Transportation Model a comparison of destination choice model and gravity model is presented in this paper. The objective is to
 augment the planning process and assist practitioners to achieve a better understanding of these two
 alternative modeling methods and make their own decision for trip distribution model.

The doubly-constrained gravity model is used as the basic form, and it is incorporated by the socioeconomic factors to adjust the non-linear friction factor calibrated. The multinomial logit destination choice model is based upon utility maximization principle, and it is identified by purpose and household stratified by income level and makes use of mode choice logsums, distance terms, zonal employment, household characteristics and regional geographic characteristics. So market segments are differentiated in both models. And parameters are calibrated to match trip length distribution of specific market segment. The results are compared to the Household Travel Survey data.

From the result, it is found that the destination choice model better replicates the observed trip length and origin-destination matrix for Home Based Work trip. This research support the hypothesis that destination choice model provides better result than gravity model in a statewide travel demand model. A suggestion concluded from this paper is to compare the outputs from two models with the survey data and choose the better one for each trip purpose.

The research suggests that though there are differences in the results between two models and destination choice model performs much better in terms of trip length and flows. In addition, the destination choice model has the advantage that it allows accounting for geographic barriers, such as psychological effects of choosing destinations involving crossing a larger river. Because of the robustness, destination choice model should be the preferred method in trip distribution models. It will effectively promote the application of the destination model in large-scale trip distribution models.

# 2 ACKNOWLEDGEMENT

3 This research is a task from a project titled "Maryland Statewide Transportation Model (MSTM)" 4 supported by Maryland State Highway Administration (SHA). MSTM is a trip based travel demand model first of its kind in the state of Maryland developed by National Center for Smart Growth Research 5 6 and education at University of Maryland and Parsons Brinkerhoff Inc. for SHA. MSTM is a continuous 7 model development process for the last four years and the model is used by a number of local, county, MPO, and at neighboring state agencies. Special thanks go to Rosella Picado and Erin Wardell for 8 implementing and calibrating the destination choice model in MSTM and appreciation to Rick Donnelley 9 and Tara Weidner for sharing their experiences on destination choice modeling (all from Parsons 10 Brinckerhoff). Useful comments by a number of personnel at NCSG and SHA are greatly appreciated. 11 The opinions, and viewpoints and expressed in this paper are solely those of the authors and does not 12 13 relate to the aforementioned agencies.

# 2 **REFERENCE**

- 3 (1) De Grange, L., E. Fernández, and J. de Cea. A Consolidated Model of Trip
- Distribution. *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 1, 2010, pp. 61–75.
- 6 (2) Hensher, D. A. Handbook of Transport Modelling. vol. 1. Emerald Group Pub Ltd, 2000.
- 7 (3) De Dios Ortúzar, J., L. G. Willumsen. Modelling Transport. Wiley, 2001.
- 8 (4) Cascetta, E., F. Pagliara, and A. Papola. Alternative Approaches to Trip Distribution Modelling: A
   9 Retrospective Review and Suggestions for Combining Different Approaches. *Papers in Regional* 10 Science, vol. 86, no. 4, 2007, pp. 597–620.
- 11 (5) De Dios Ortúzar, J., L. G. Willumsen. *Modelling Transport*. Wiley, 2011.
- Evans, A. W. Some Properties of Trip Distribution Methods. *Transportation Research*, vol. 4, no. 1, 1970, pp. 19–36.
- (7) Casey, H. J.Applications to Traffic Engineering of the Law of Retail Gravitation. *Traffic Quarterly*, vol. 9, no. 1, 1955, pp. 23–25.
- 16 (8) Wilson, A.A Statistical Theory of Spatial Distribution Models. 1967.
- (9) Anderson, J. E.A Theoretical Foundation for the Gravity Rquation. *The American Economic Review*,
   vol. 69, no. 1, 1979, pp. 106–116.
- (10) Haynes, K. E., and A. S. Fotheringham. *Gravity and Spatial Interaction Models*. Sage Publications
   Beverly Hills, 1984.
- (11) Erlander,S., and N. F. Stewart, *The Gravity Model in Transportation Analysis: Theory and Extensions*, vol. 3. Vsp, 1990.
- (12) Cascetta, E., and A. Papola. A Trip Distribution Model Involving Spatial and Dominance
   Attributes. *Computer-Aided Civil and Infrastructure Engineering*, vol. 23, no. 2, 2008, pp. 116–124.
- (13) Fotheringham, A. S.Some Theoretical Aspects of Destination Choice and Their Relevance to
   Production-constrained Gravity Models. *Environment and Planning A*, vol. 15, no. 8, 1983, pp. 1121–1132.
- (14) Fotheringham, A. S.A New Set of Spatial-interaction Models: the Theory of Competing
   Destinations. *Environment and Planning A*, vol. 15, no. 1, pp. 15–36, 1983.
- (15) Fotheringham, A. S. Modelling Hierarchical Destination Choice. *Environment and planning A*, vol. 18, no. 3, 1986, pp. 401–418.
- (16) M. E. Ben-Akiva and S. R. Lerman, *Discrete Choice Analysis: Rheory and Application to Travel Demand*, vol. 9. The MIT Press, 1985.
- (17) Woodside, A. G., and S. Lysonski. A General Model of Traveler Destination Choice. *Journal of Travel Research*, vol. 27, no. 4, 1989, pp. 8–14.
- (18) Swait, J., and M. Ben-Akiva. Empirical Test of a Constrained Choice Discrete Model: Mode Choice
   in Sao Paulo, Brazil. *Transportation Research Part B: Methodological*, vol. 21, no. 2, 1987, pp.
   103–115.
- (19) Cascetta, E., and A. Papola. Random Utility Models with Implicit Availability/Perception of Choice
   Alternatives for the Simulation of Travel Demand. *Transportation Research Part C: Emerging Technologies*, vol. 9, no. 4, 2001, pp. 249–263.
- 42 (20) Seddighi,H. R., and A. L. Theocharous. A Model of Tourism Destination Choice: a Theoretical and
   43 Empirical Analysis. *Tourism Management*, vol. 23, no. 5, 2002, pp. 475–487.
- (21) Louviere, J. J. Random Utility Theory-based Stated Preference Elicitation Methods: Applications in
   Health Economics with Special Reference to Combining Sources of Preference Data. 2004.
- 46 (22) Zheng, J., and J. Y. Guo. Destination Choice Model Incorporating Choice Set Formation. In
- 47 *Transportation Research Record: Journal of the Transportation Research Board*, No. 2690,
- 48 Transportation Research Board of National Academies, Washington, D.C.,2008.

- (23) Murchland, J. D., L. S. of Economics, and P. S. T. N. T. Unit. *Some Remarks on the Gravity Model of Traffic Distribution, and an Equivalent Maximization Formulation*. London School of Economics
   and Political Science, Transport Network Theory Unit, 1966.
- 4 (24) Evans, S. P., and H. R. Kirby. A Three-dimensional Furness Procedure for Calibrating Gravity
   5 Models. *Transportation Research*, vol. 8, no. 2, 1974, pp. 105–122.
- 6 (25) Isard,W.A Simple Rationale for Gravity Model Type Behavior.*Papers in Regional Science*, vol. 35, no. 1, 1975, pp. 25–30.
- 8 (26) Mikkonen,K., and M. Luoma. The Parameters of the Gravity Model are Changing-How and
   9 Why? *Journal of Transport Geography*, vol. 7, no. 4, 1999, pp. 277–283.
- 10 (27) Metropolitan Travel Forecasting: Current Practice and Future Direction, vol. 288. Natl Academy
   Pr, 2007.
- (28) Thill,J. C.Choice Set Formation for Destination Choice Modelling.*Progress in Human Geography*,
   vol. 16, no. 3, 1992, pp. 361–382.
- (29) Thill,J. C., and J. L. Horowitz. Travel-Time Constraints on Destination-Choice Sets. *Geographical Analysis*, vol. 29, no. 2, 1997, pp. 108–123.
- (30) Bhat,C., A. Govindarajan, and V. Pulugurta. Disaggregate Attraction-end Choice Modeling:
   Formulation and Empirical Analysis.In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1645, Transportation Research Board of National Academies,
   Washington, D.C., 1998, pp. 60–68.
- (31) Pozsgay,M. A., and C. R. Bhat. Destination Choice Modeling for Home-based Recreational Trips:
   Analysis and Implications for Land Use, Transportation, and Air Quality Planning.In
   *Transportation Research Record: Journal of the Transportation Research Board*, No. 1777,
   Transportation Research Board of National Academies, Washington, D.C.,2001,pp. 47–54.
- 24 (32) Train, K., Discrete Choice Methods with Simulation. Cambridge Univ Pr, 2003.
- (33) Vrtic, M., P. Fröhlich, N. Schüssler, K. W. Axhausen, D. Lohse, C. Schiller, and H. Teichert. Twodimensionally Constrained Disaggregate Trip Generation, Distribution and Mode Choice Model:
  Theory and Application for a Swiss National Model. *Transportation Research Part A: Policy and Practice*, vol. 41, no. 9, 2007, pp. 857–873.
- (34) Russo,F., and A. Comi. A Modelling System to Link End-Consumers and Distribution Logistics.
   2010.
- (35) Ye,X., W. Cheng, and X. Jia. A Synthetic Environment to Evaluate Alternative Trip Distribution
   Models.In *Transportation Research Record: Journal of the Transportation Research Board*, No.
   2784, Transportation Research Board of National Academies, Washington, D.C.,2012.
- (36) Duffus, L. N., A. Sule Alfa, and A. H. Soliman. The Reliability of Using the Gravity Model for
   Forecasting Trip Distribution. *Transportation*, vol. 14, no. 3, 1987, pp. 175–192.
- (37) Levinson, D., and A. Kumar. A Multi-modal Trip Distribution Model.In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1466, Transportation Research Board
   of National Academies, Washington, D.C., 1994, pp. 124-131.
- (38) Kim,C., C. G. Choi, S. Cho, and D. Kim. A Comparative Study of Aggregate and Disaggregate
   Gravity Models Using Seoul Metropolitan Subway Trip Data. *Transportation Planning and Technology*, vol. 32, no. 1, 2009, pp. 59–70.
- 42 (39) Bernardin,L.,V., F. Koppelman, and D. Boyce. Enhanced Destination Choice Models Incorporating
   43 Agglomeration Related to Trip Chaining while Controlling for Spatial Competition.In
   44 *Transportation Research Record: Journal of the Transportation Research Board*, No. 2132,
   45 Transportation Research Board of National Academies, Washington, D.C.,2009,pp. 143–151.
- (40) Chow, L. F., F. Zhao, M. T. Li, and S. C. Li. Development and Evaluation of Aggregate Destination
   Choice Models for Trip Distribution in Florida. In *Transportation Research Record: Journal of the*
- *Transportation Research Board*, No. 1931, Transportation Research Board of National Academies,
   Washington, D.C.,2005,pp. 18–27.
- 50 (41) Bureau of Economic Analysis: Bureau of Economic Analysis: SA25N Total full-time and part-time
   51 employment by NAICS industryURL:http://www.bea.gov/, Date Accessed, July, 2012.

Mishra et al.

1 (42) Census Summary Files, URL: http://www.census.gov/, Date Accessed, July, 2012.