ONLINE SUPPLEMENT

PART 1: EXTENDED BASS MODEL

The Bass model is derived from a hazard function and depicts the probability of adoption at time t given that it has not yet occurred as $\frac{f(t)}{1-F(t)} = p + qF(t)$ in which $f(t)$ is the probability of adoption and $F(t)$ is the cumulative probability of adoption until time t. The coefficients of innovation and emulation are indicated by p and q respectively. If m is the size of potential market, then Bass model can be rewritten as $\frac{dY(t)}{dt} = y(t) = (pm + qY(t)) \times (1 - F(t))$ in which $y(t)$ is the number of adoptions in time t and $Y(t)$ is the total number of adoptions until time t.

Let the probability density of adoption in group i at time t be $f_i(t)$. We extend the traditional Bass model to simultaneous adoptions in k separate but inter-linked groups as follows

$$f_i(t) = [p_i + \sum_{j=1}^{k} q_{ji} F_j(t)][1 - F_i(t)] \quad (1)$$

Where q_{ji} captures the network effects. When $i = j$, q_{ji} is simplified as q_{ii} and captures the direct network effect within group i. When $i \neq j$, q_{ji} captures the indirect network effect from group j on group i. Note that when $n = 1$, the above model is reduced to the original Bass single-group linear model. Due to the fact that $\frac{dF_i(t)}{dt} = f_i(t)$ we can consider equation (1) as a first order differential equation and solve for $F_i(t)$ to derive a set of k simultaneous equations.

By (1), $\frac{dF_i(t)}{dt} = [p_i + \sum_{j=1}^{k} q_{ji} F_j(t)][1 - F_i(t)]$, which is

$$\int \frac{1}{[p_i + \sum_{j=1}^{k} q_{ji} F_j(t)][1 - F_i(t)]} dF_i(t) = t + D, \text{ or}$$

$$\int \frac{A - AF_i(t) + Bp_i + B\sum_{j=1}^{k} q_{ji} F_j(t)}{[p_i + \sum_{j=1}^{k} q_{ji} F_j(t)][1 - F_i(t)]} dF_i(t) = t + D.$$

$$A + Bp_i + B \sum_{j=1, j\neq i}^{k} q_{ji} F_j(t) = 1 \text{ and } -A + Bq_{ii} = 0.$$

$A = Bq_{ii}$ and $Bq_{ii} + Bp_i + B\sum_{j=1, j\neq i}^{k} q_{ji}F_j(t) = 1$.

$$A = \frac{q_{ii}}{p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t)} \text{ and } B = \frac{1}{p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t)}.$$

$$\int [\frac{\frac{q_{ii}}{p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t)}}{p_i + \sum_{j=1}^{k} q_{ji}F_j(t)} + \frac{\frac{1}{p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t)}}{1 - F_i(t)}] dF_i(t) = t + D$$ which can be simplified as

$$\frac{q_{ii}(1 - F_i(t))}{p_i + \sum_{j=1}^{k} q_{ji}F_j(t)} = e^{-(t+D)(p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t))}$$ we can derive:

$$F_i(t) = \frac{q_{ii} - (p_i + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t))e^{-(t+D)(p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t))}}{q_{ii}(e^{-(t+D)(p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t))} + 1)}.$$

Since $F_j(0) = 0$, $j = 1, ..., k$, we have

$$\frac{q_{ii} - p_i e^{-D(p_i + q_{ii})}}{q_{ii}(e^{-D(p_i + q_{ii})} + 1)} = 0, \text{ or}$$

$$-D = \frac{1}{p_i + q_{ii}} \ln \frac{q_{ii}}{p_i}.$$

Then

$$F_i(t) = \frac{1 - (\frac{p_i}{q_{ii}} + \frac{1}{q_{ii}} \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t))e^{-(t - \frac{1}{p_i + q_{ii}} \ln \frac{q_{ii}}{p_i})(p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t))}}{e^{-(t - \frac{1}{p_i + q_{ii}} \ln \frac{q_{ii}}{p_i})(p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t))} + 1},$$

where $e^{-(t - \frac{1}{p_i + q_{ii}} \ln \frac{q_{ii}}{p_i})(p_i + q_{ii} + \sum_{j=1, j\neq i}^{k} q_{ji}F_j(t))}$ can be simplified as

$$e^{-(p_i+q_{ii}+\sum_{j=1,j\neq i}^{k} q_{ji}F_j(t))t+(p_i+q_{ii})\frac{1}{p_i+q_{ii}}\ln\frac{q_{ii}}{p_i}+\sum_{j=1,j\neq i}^{k} q_{ji}F_j(t)\frac{1}{p_i+q_{ii}}\ln\frac{q_{ii}}{p_i}}$$

$$=\frac{q_{ii}}{p_i}e^{-(p_i+q_{ii})t+\sum_{j=1,j\neq i}^{k} q_{ji}F_j(t)(\frac{1}{p_i+q_{ii}}\ln\frac{q_{ii}}{p_i}-t)}.$$

Then

$$F_i(t)=\frac{1-(1+\frac{1}{p_i}\sum_{j=1,j\neq i}^{k} q_{ji}F_j(t))e^{-(p_i+q_{ii})t+\sum_{j=1,j\neq i}^{k} q_{ji}F_j(t)(\frac{1}{p_i+q_{ii}}\ln\frac{q_{ii}}{p_i}-t)}}{\frac{q_{ii}}{p_i}e^{-(p_i+q_{ii})t+\sum_{j=1,j\neq i}^{k} q_{ji}F_j(t)(\frac{1}{p_i+q_{ii}}\ln\frac{q_{ii}}{p_i}-t)}+1}.$$

Which results in:

$$Y_i(t)=m_iF_i(t)=m_i\frac{1-(1+\frac{1}{p_i}\sum_{j=1,j\neq i}^{k} q_{ji}F_j(t))e^{-(p_i+q_{ii})t+\sum_{j=1,j\neq i}^{k} q_{ji}F_j(t)(\frac{1}{p_i+q_{ii}}\ln\frac{q_{ii}}{p_i}-t)}}{\frac{q_{ii}}{p_i}e^{-(p_i+q_{ii})t+\sum_{j=1,j\neq i}^{k} q_{ji}F_j(t)(\frac{1}{p_i+q_{ii}}\ln\frac{q_{ii}}{p_i}-t)}+1} \quad (2)$$

In equation (2), $Y_i(t)$ is the total number of adaptors in group i at time t. It is affected by both external sources, p_i, and the number of adaptors until time t in not only group i but also other groups $(j \neq i)$, and thus captures both direct and indirect network effects

PART 2. DETAILS OF DATA SETS AND METHODS

Multi-dimensional Scaling and Ward Clustering

The patient flow may be affected by organizational factors such as similar affiliations of the doctors. However, the decision to refer a patient is made in two stages: first, a doctor decides about the specialty to which a patient should be referred based on the patient's medical condition; next, he decides about the doctor with that specific specialty to whom the patient should be referred. Choosing the "specialty" for referring a patient is based on patients' medical necessities and needs. Physicians' affiliations to common practices only affect the second stage of patient referral decisions. We group "Specialties" rather than individual physicians, which in part eliminates the possible confounding factors such as similar affiliations. Furthermore, our approach of clustering by specialties rather than clustering by practices results in generalizable insights about the effects of specialties on each other in HIE adoption. These clusters and their interlinked effects can easily be used by RHIOs and other HIE affiliated organizations to promote HIE through targeted marketing efforts.

Patient flows among specialties may be driven by a variety of factors such as rural/urban divide, affiliations with common hospitals, specialty size, and closeness of the medical specialties themselves such as allergy and dermatology for instance. Please note that our objective is not to derive any relationships between these various factors and adoption; we focus only on patient flows and adoption. However, there may be a myriad of such factors underlying patient flows.

In order to identify the clusters of specialists with highest ratio of common patients between them, we first construct a matrix of specialties in which each element of matrix shows the ratio of common patients between the two specialties. Specialties with high ratio of common patients are the ones that have the highest flow of patients between each other. Based on this matrix, we applied Multi-dimensional Scaling (MDS) to create three artificial dimensions for each specialty such that the artificial distance between the two specialties based on these dimensions are correlated as highly as possible with the real common patients' ratio between them.[51,52] These dimensions were then used to construct 3 clusters of specialties. We used the Ward minimum-variance clustering method. The minimum variance method is designed to generate clusters in such a way as to minimize the within cluster variance. In this method, the distance between two clusters is the ANOVA sum of squares between the two clusters added up over all the variables. At each generation, the within-cluster sum of squares is minimized over all partitions obtainable by merging two clusters from the previous generations.[53–55]

Figure 3 shows the relative position of 21 specialties in an artificial three dimensional space. The closer the specialties are, the higher is the common patient ratio between them. Table 3 presents the complete list of specialties within each cluster. We used PROC MDS and PROC CLUSTER in SAS to conduct MDS analysis and Cluster specialties accordingly.[54] For a complete review of the methods applied on this data set see Yaraghi et al.[41]

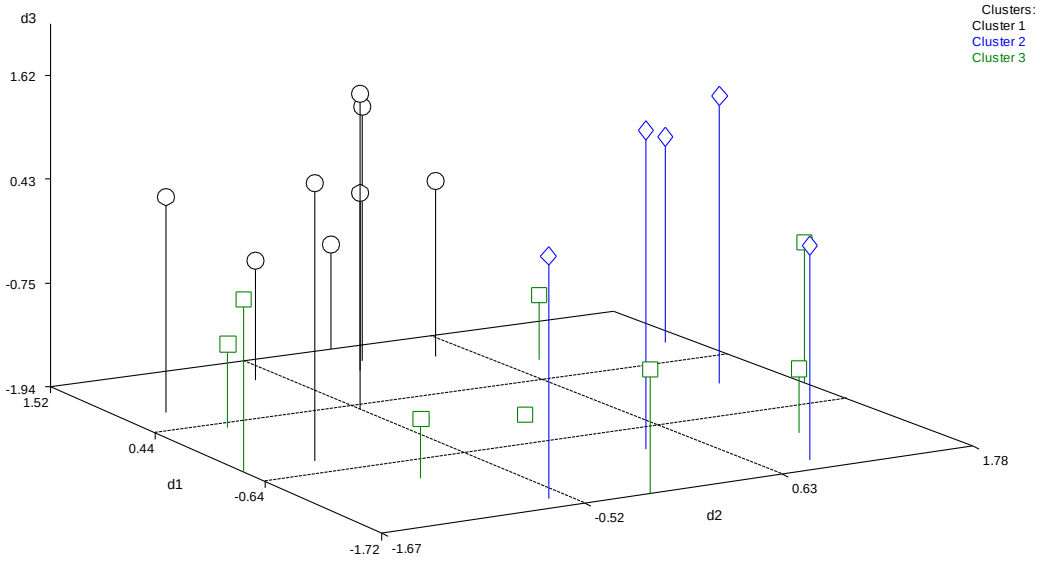

Figure 3. Cluster of specialties based on common patients

We also clustered practices based on the two dimensions of latitude and longitude. Since we already have two real dimensions for clustering, we do not need to produce any more dimensions by MDS analysis and can directly apply clustering methods. Figure 4 shows the 2 clusters of practices based on their geographical location. The practices shown with black stars are the ones which are concentrated in urban areas of western New York and while the ones shown with blue stars are the practices which are farther from each other and are located in rural areas.

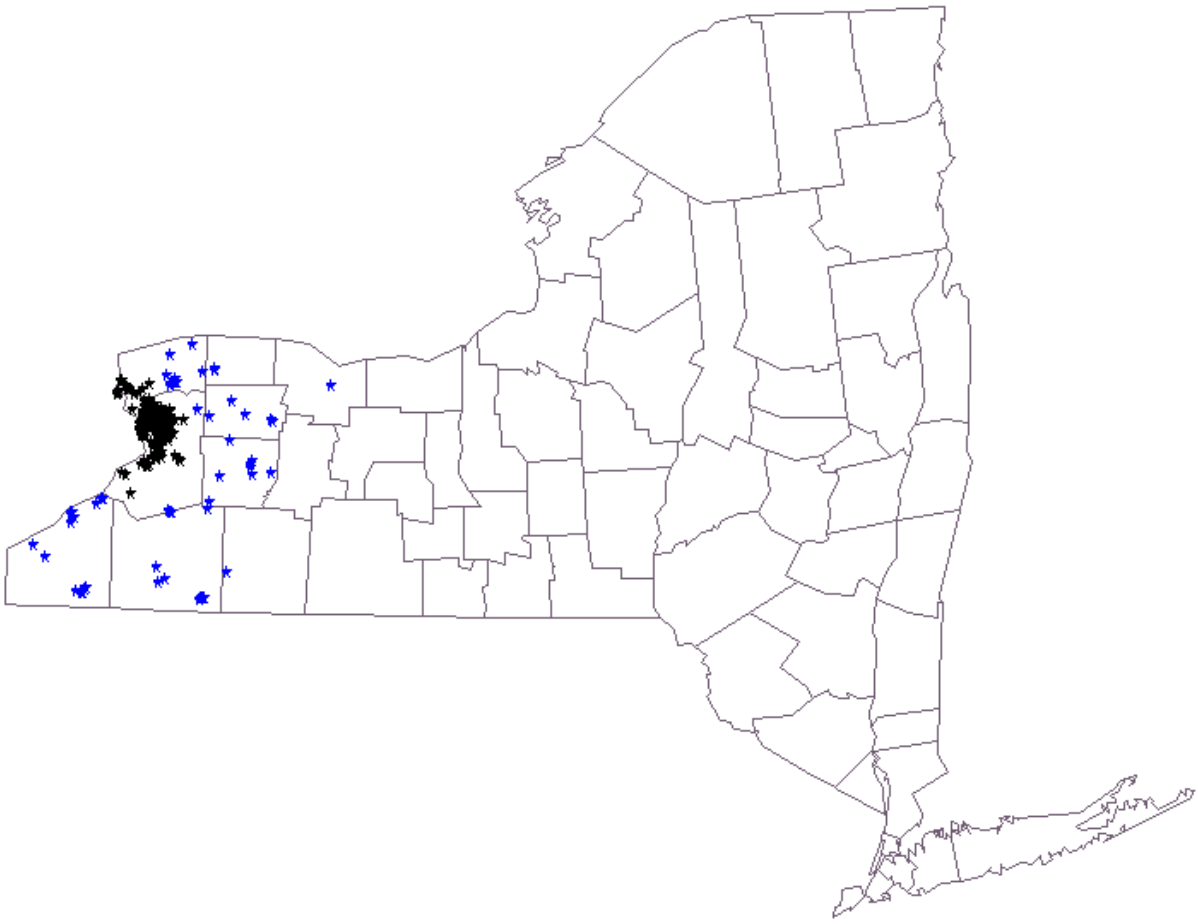

Figure 4. clusters of practices based on their latitude and longitude

Nonlinear three stage least squares method

After identifying different clusters, the sets of equations in (2) are estimated jointly using the nonlinear three-stage least squares option of MODEL procedure in SAS software. For studying the effect of specialty groups, equations in (2) will be a set of 3 different simultaneous equations $(i = 1,2,3)$ while for examining the effects of geographical location, since we have two groups of urban and rural locations, equations in (2) will be a set of 2 different simulations equations$(i = 1,2)$. The ordinary least squares (OLS) estimation method is not appropriate because the estimators of the structural coefficients

are biased and inconsistent due to the simultaneity bias. Instead, the methods of two-stage least squares (2SLS) or three-stage least squares (3SLS) should be used. Three stage least squares method was suggested by Zellner & Theil[56] to estimate simultaneous equation systems. It is a combination of two stage least squares and seemingly unrelated regression methods and uses the two-stage least squares estimated moment matrix of the structural disturbances to estimate all coefficients of the entire system simultaneously. The major difference between 2SLS and 3SLS lies in the assumptions underlying the random errors in the simultaneous equations. If the random errors are correlated, then 3SLS is more appropriate than 2SLS because it produces more efficient estimates. Such correlations among the random errors could be present if other possible contingency variables are unintentionally omitted from the simultaneous contingency model, leaving the influence of these omitted variables to be absorbed by the random errors of the equations and consequently, rendering the random errors.[57] Nonlinear least squares method has been shown to produce more efficient estimates in the context of the Bass diffusion model.[58]

	Cluster	Cluster Name	Specialty
1	1	medical and surgical subspecialties (especially cardiac)	INTERNAL MEDICINE - CARDIOVASCULAR DISEASE
2			INTERNAL MEDICINE - ENDOCRINOLOGY DIABETES & METABOLISM
3			INTERNAL MEDICINE – GASTROENTEROLOGY
4			INTERNAL MEDICINE - HEMATOLOGY & ONCOLOGY
5			OTOLARYNGOLOGY
6			PEDIATRICS
7			RADIOLOGY
8			SURGERY
9	2	primary care and women's health	ADULT MEDICINE
10			ALLERGY & IMMUNOLOGY – ALLERGY
11			FAMILY MEDICINE
12			INTERNAL MEDICINE - INFECTIOUS DISEASE
13			WOMEN'S HEALTH
14	3	medical and surgical subspecialties (especially renal, neurological, and musculoskeletal)	INTERNAL MEDICINE - GERIATRIC MEDICINE
15			INTERNAL MEDICINE – NEPHROLOGY
16			INTERNAL MEDICINE - PULMONARY DISEAS
17			INTERNAL MEDICINE – RHEUMATOLOGY
18			ORTHOPEDICS
19			PSYCHIATRY & NEUROLOGY – NEUROLOGY
20			PSYCHIATRY & NEUROLOGY – PSYCHIATRY
21			UROLOGY

Table 3: Cluster Members

PART 3: ANALYSIS OF ADOPTION AT PRACTICE LEVEL

We have also analyzed the effects of patient flow on HIE adoption at practice level. We have followed a very similar clustering approach as described in previous section to group practices into 3 different clusters based on their common patients. In each cluster, the ratio of common patients between practices is high while the ratio of common patients between practices in two different clusters is low. Following the same argument about the strong effects of patient flow on deriving HIE adoption, we expect to see strong within group effects and weak between group effects in HIE adoption at practice level.

Parameter	*Description*	*Estimate*	*Std. Err.*	*t-value*	$P_r > \|t\|$
q_{11}	emulation effect within group 1	0.064273	0.0215	2.98	0.0061
q_{22}	emulation effect within group 2	0.05915	0.0261	2.26	0.0321
q_{33}	emulation effect within group 3	0.056531	0.0163	3.48	0.0018
p_1	innovation effect in group 1	0.025945	0.000966	26.87	<.0001
p_2	innovation effect in group 2	0.015332	0.00573	2.67	0.0128
p_3	innovation effect in group 3	0.021859	0.0114	1.92	0.0654
q_{12}	emulation effect from group 1 on group 2	0.009138	0.00316	2.89	0.0077
q_{13}	emulation effect from group 1 on group 3	0.009636	0.00165	5.85	<.0001
q_{21}	emulation effect from group 2 on group 1	0.017105	0.00906	1.89	0.0702
q_{23}	emulation effect from group 2 on group 3	0.042804	0.0837	0.51	0.6132
q_{31}	emulation effect from group 3 on group 1	0.034059	0.0377	0.90	0.3748
q_{32}	emulation effect from group 3 on group 2	0.017038	0.00499	3.41	0.0021

Table 4: The estimation of HIE adoption model at practice level

Table 4 shows the results of the model estimation at practice level. The within group emulation effects shown by q_{11}, q_{22} and q_{33} are all positive and significant at 5% level. Some of the between group effects are also significant. However, the magnitude of the between group effects are much less than the within group effects. These results are consistent with our main findings and confirms positive effects of common patients on HIE adoption.

REFERENCES

1. Schiffman SS, Reynolds ML, Young FW, Carroll JD. *Introduction to multidimensional scaling: Theory, methods, and applications*. Academic Press New York; 1981. Available at: http://www.getcited.org/pub/102136264. Accessed August 8, 2012.

2. Kruskal JB. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*. 1964;29(1):1–27.

3. Punj G, Stewart DW. Cluster analysis in marketing research: Review and suggestions for application. *Journal of marketing research*. 1983;20(2):134–148.

4. SAS. *SAS User's Guide (Release 9.2)*. SAS Institute, Inc; 2004.

5. Ward Jr JH. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*. 1963;58(301):236–244.

6. Yaraghi N, Du AY, Sharman R, Gopal R, Ramesh R. Clustering Physicians Based on Professional Proximity Using SAS. In: *SAS Global Forum 2012 conference*. SAS Institute, Inc; 2012.

7. Zellner A, Theil H. Three-stage least squares: simultaneous estimation of simultaneous equations. *Econometrica: Journal of the Econometric Society*. 1962;30(1):54–78.

8. Lin WT, Shao B. The relationship between user participation and system success: a simultaneous contingency approach. *Information & Management*. 2000;37(6):283–295.

9. Srinivasan V, Mason CH. Technical Note—Nonlinear Least Squares Estimation of New Product Diffusion Models. *Marketing science*. 1986;5(2):169–178.