Building a composite Help-Wanted Index over 1951-2009

Barnichon, Regis

Federal Reserve Board

15 December 2009

Building a composite Help-Wanted Index over 1951-2009*

Regis Barnichon
Federal Reserve Board

15 December 2009
(first draft November 2009)

Abstract

This paper presents a measure of vacancy posting that captures the behavior of total –print and online– help-wanted advertising. By modeling the share of online job advertising as the diffusion of a new technology –online job posting and job search– I can combine information on both print and online help-wanted advertising. I use this consistent measure of vacancy posting over 1951-2009 to estimate the elasticity of the matching function and find an estimate of 0.58.

*Comments welcome. I would like to thank Bruce Fallick, Chris Nekarda and John M. Roberts for helpful comments. The views expressed here do not necessarily reflect those of the Federal Reserve Board or of the Federal Reserve System. Any errors are my own. E-mail: regis.barnichon@frb.gov

1 Introduction

The number of job openings, or vacancies, is an important indicator of the state of the labor market and is used extensively in policy and academic circles.[1]

The traditional measure of vacancy posting is the Conference Board Help-Wanted Index (HWI) that measures the number of help-wanted advertisements in 51 major newspapers. However, since the mid-1990s, this "print" measure of vacancy posting has become increasingly unrepresentative as advertising over the internet has become more prevalent. In fact, the Conference Board stopped publishing its print HWI in May 2008, and since 2005, publishes instead a measure of "online" help wanted advertising.

In this paper, I build a vacancy posting index over 1951-2009 that captures the behavior of total –"print" and "online"– help-wanted advertising. A key aspect of this exercise is to estimate the share of newspaper help-wanted advertising in total advertising so that one may combine information from both print and online indexes. Since the print share is not directly observable, I estimate it in two steps.

First, I combine different datasets to estimate the empirical behavior of the newspaper share over 2000-2009. In particular, since the print HWI and the online HWI are simultaneously observable since May 2005, I can recover the newspaper share over that period. Further, by assuming that the true composite HWI comoves with the Job Openings and Labor Turnover Survey (JOLTS) measure of job openings, I can recover a second estimate of the newspaper share over 2000-2009. The behavior of the inferred print shares suggests that the share of online job advertising follows an S-curve, a shape characteristic of the process of technology diffusion (see e.g. Comin, Hobijn and Rovito, 2007).

Second, to recover the share of print advertising since the mid-1990s, I model the development of online job advertising as the diffusion of a new technology: online job posting and job search. To choose an appropriate functional form for the diffusion of online job advertising, I use data from the World Development Indicators on the percentage of internet users in the US population since 1990. I assume that the diffusion of online job posting follows the same pattern as the diffusion of the internet. The best fit is given by a Mixed-Information Source Model; a model widely used in the marketing literature (see e.g. Geroski, 2000). I then estimate the share of print advertising by fitting that model to the print share implied by JOLTS data over 2000-2009 and by imposing a share of one in 1994, which roughly corresponds to the introduction of the World Wide Web.

To illustrate this new composite Help-Wanted index, I estimate a matching function over

[1] See Abraham and Katz (1986), Abraham (1987), Blanchard and Diamond (1989) and more recently Shimer (2005).

1951-2009 using Shimer's (2007) measure of the job finding rate as the dependent variable. When imposing a constant returns to scale specification, I find an estimate of 0.58, a value consistent with previous estimates (see Petrongolo and Pissarides, 2001). However, after allowing for different elasticities with respect to vacancies and unemployment, I find evidence of mildly decreasing returns to scale with an estimated degree of returns to scale of around 0.80.

This paper is related to the recent composite HWI index built by Fallick (2008). Fallick estimates the newspaper share from a low-frequency trend in print HWI since 1994. However, fitting a low-frequency trend to the raw print HWI series will estimate the share of print advertising consistently only if the decline in that share is "slow" enough to be interpreted as a low frequency movement. However, during the fast diffusion stage of online advertising, there is no guarantee that this is the case. As a result, a low-frequency filter applied to the print HWI may not be able to disentangle the diffusion of online advertising from the cyclical drop in newspaper advertising in the 2001 and 2008 recessions. In fact, I show that, after 2008, the print share recovered with the low-frequency trend in print HWI strongly underestimates the print share directly inferred from the print and online HWIs, and misses the inflection in the diffusion rate of online advertising.

This paper is organized as follows. Section 1 briefly describes the construction of a composite index under the assumption that an estimate of the share of print advertising is known; Section 2 presents a method to produce an estimate of that share; Section 3 fits a matching function to the data; and Section 4 offers some concluding remarks.

2 Constructing a composite Help-Wanted index from the share of newspaper help-wanted advertising

Denote respectively $P_t^{\#}$ and $O_t^{\#}$ the number of print help-wanted advertisements and online help-wanted advertisements. The total number of advertisements (print and online) $H_t^{\#}$ is then given by $H_t^{\#} = P_t^{\#} + O_t^{\#}$, and $s_t^p = \frac{P_t^{\#}}{P_t^{\#}+O_t^{\#}}$ is the share of print help-wanted advertising in total advertising. Further, denote H_t the composite Help-Wanted advertising index, P_t the print index, and O_t the online index. The indexes are defined with respect to some base year t_0, and I have $P_t = \frac{P_t^{\#}}{P_{t_0}^{\#}}$ and $O_t = \frac{O_t^{\#}}{O_{t_0}^{\#}}$.

To build the composite index, I consider three separate periods:

1) Before 1995:

As in Fallick (2008), I assume that there is no online job posting up until 1995, which corresponds to the introduction of the World Wide Web. As a result, $H_t^{\#} = P_t^{\#}$ and I normalize the composite index so that $H_t = P_t$ over 1951-1994.

2) January 1995-May 2005:

Over that period, the print HWI is observable but the online HWI is not. However, with the share of print advertising over that period, I can easily recover $O_t^{\#}$ from the definition of s_t^p with $O_t^{\#} = P_t^{\#}\frac{1-s_t^p}{s_t^p}$. Since the total number of job advertisements is $H_t^{\#} = P_t^{\#} + P_t^{\#}\frac{1-s_t^p}{s_t^p} = \frac{P_t^{\#}}{s_t^p}$, starting in January 1995, I construct the composite index from $d\ln H_t = d\ln H_t^{\#} = d\ln\frac{P_t^{\#}}{s_t^p} = d\ln\frac{P_t}{s_t^p}$.

3) After June 2005:

Over that (short) period, both the print and online HWI are available, and I can combine the two series.[2] Log-linearizing $H_t^{\#} = P_t^{\#} + O_t^{\#}$, I get $d\ln H_t^{\#} = s_{t-1}^p d\ln P_t^{\#} + (1-s_{t-1}^P) d\ln O_t^{\#}$ to a first order. Starting in June 2005, I can then construct the composite index from $d\ln H_t = s_{t-1}^p d\ln P_t + (1-s_{t-1}^p) d\ln O_t$.

3 Estimating the share of newspaper help-wanted advertising over 1995-2005

An important aspect of constructing a composite HWI is thus to estimate the share of newspaper advertising over 1995-2005.

As a starting point, I proceed as in Fallick (2008), and I interpret the downward trend in print HWI (see Figure 1) over 1995-2009 as a secular decline in the share of newspaper advertising due to the emergence of online advertising and the world wide web. I also assume that there are no cyclical fluctuations in the ratio of print advertising to online advertising. In that case, the behavior of is entirely explained by the downward trend in print advertising. With those two assumptions, I can indirectly estimate the share of print advertising: I fit a quartic polynomial to print HWI over 1951-2009, and I estimate the print share at time t as the ratio of the polynomial's value at time t to the polynomial's value in 1994. Figure 2 plots the estimated print share.

In this section, I present a simple way to directly recover the behavior of the newspaper share over 2000-2009 by combining information from different indexes. The behavior of the print share over 2000-2009 suggests that fitting a low-frequency trend directly to the print HWI does not correctly capture the behavior of the print share over 1995-2009. Moreover, the print share appears to follow an S-curve, a shape characteristic of the process of technology diffusion (see e.g. Comin, Hobijn and Rovito, 2007). Accordingly, I model the development of online job advertising as the diffusion of a new technology: online job posting and job search

[2] While the Conference Board does not publish its print HWI since June 2008, it still produces the index.

3.1 Inferring the print share from the print and online HWIs and JOLTS data

A simple way to directly recover the behavior of the print share over 2005-2009 is to use the fact that print advertising and online advertising are simultaneously available over that period. Indeed, from the definition of s_t^p, I have $O_t^{\#} = P_t^{\#}\frac{1-s_t^p}{s_t^p}$ so that after log-differencing, I can infer the behavior of $\hat{s}_t^{p,HWI}$, the share of print advertising, from

$$d\ln\left(\frac{s_t^p}{1-s_t^p}\right) = d\ln P_t - d\ln O_t. \tag{1}$$

Further, I can use JOLTS data on the number of job openings, J_t, to recover the behavior of the print share over 2000-2009.[3] I assume that the true composite index approximately comoves with the JOLTS series, so that I can write

$$\ln H_t = \ln\alpha + \ln J_t + \varepsilon_t \text{ with } \varepsilon_t \ll 1 \tag{2}$$

with $\ln\alpha$ a constant. Importantly, I will be able to verify the assumption $\varepsilon_t \ll 1$ by using the estimated share $\hat{s}_t^{p,HWI}$ as a reference point. Using the fact that $P_t^{\#} = s_t H_t^{\#}$ and differencing (2), I can infer a value $\hat{s}_t^{p,JOLTS}$ of the share of print advertising from

$$d\ln s_t^p = d\ln P_t - d\ln(J_t). \tag{3}$$

Figure 2 plots $\hat{s}_t^{p,HWI}$ and $\hat{s}_t^{p,JOLTS}$along with the print share estimated from the polynomial trend.

Two observations are worth noting:

First, we can see that $\hat{s}_t^{p,JOLTS}$ tracks $\hat{s}_t^{p,HWI}$ remarkably well over 2005-2009. This confirms my initial assumption that the true composite HWI approximately comoves with the JOLTS series. Thus, using JOLTS data can provide valuable information on the behavior of the print share over 2000-2005.

Second, the rate of decline of the estimated print shares is not constant and is slower before 2004 and after 2008. In contrast, the polynomial trend predicts an accelerating rate of decline and strongly overestimates the share of print advertising after 2008. Going forward, the polynomial trend predicts that the share of newspaper advertising will be essentially zero by 2011, a probably too bleak scenario for newspaper advertising.[4]

[3] JOLTS is produced by the BLS and contains monthly data on job openings from 16,000 establishments since December 2000.

[4] A low-frequency trend will estimate the share of print advertising consistently only if the decline in that share is "slow" enough to be interpreted as a low frequency movement. However, during the fast diffusion stage

The behavior of the print share over 2000-2009 suggests that fitting a low-frequency trend directly to the print HWI may not be the best approach to estimate the behavior of the print share over 1995-2009. Moreover, the print share appears to follow an S-curve, a shape characteristic of the process of technology diffusion (see e.g. Comin, Hobijn and Rovito, 2007). In the next section, I model the development of online job advertising as the diffusion of a new technology: online job posting and job search.

3.2 Modeling the diffusion of online job advertising

At least since Griliches (1957), economists have acknowledged the good approximation that the S-shaped curves such as the Logistic provide to model the process of technology diffusion.[5] However, a number of researchers, notably Dixon (1980) and more recently Comin, Hobijn and Rovito (2007), showed that the logistic function may not provide the best description of technology diffusion, in particular because the late adoption stages of a technology appear to occur a lot slower than the logistic would predict.[6]

To choose an appropriate functional form for the diffusion of online job advertising, I exploit the fact that the extent of online advertising depends directly on the number of internet users. I use data from the World Development Indicators on the percentage of internet users in the US population over 1990-2008, and I assume that the diffusion of online job posting follows the same pattern as the diffusion of the internet. Figure 3 plots the share of internet users in the US, and shows that the diffusion of the internet is indeed asymmetric with a slower diffusion rate in the later years. I compare the performance of three popular models of technology diffusion; the Logistic function, the Gompertz function, popularized in economics by Dixon (1980), and the Mixed Information Source Model (MISM), widely used in the marketing literature (Geroski, 2000).[7] The latter two functions are asymmetric and allow for an adoption rate that is much faster during the initial adoption stages than the latter stages. Compared to the Gompertz curve, the mixed information source model allows for even more asymmetry in the diffusion process. Figure 3 shows the non-linear least-square fits of these three models on the diffusion of the internet. The mixed information source model does a very good job at

of online advertising, there is no guarantee that this is the case. As a result, a low-frequency filter may not be able to disentangle the cyclical drop in newspaper advertising in the 2001 and 2008 recessions from the diffusion of online advertising.

[5] The logistic curve is defined by $X_t = \frac{1}{1+\alpha e^{-\beta t}}$, with β reflecting the speed of adoption and α a constant of integration positioning the curve on the time scale. The logistic function is a solution to the differential equation $\frac{\dot{X}_t}{X_t} = \beta(1 - X_t)$, corresponding to a diffusion process in which the source of diffusion is previous users.

[6] The logistic function is symmetric around its inflection point.

[7] The Gompertz function is $X_t = e^{\alpha e^{-\beta t}}$ and is a solution to the differential equation $\frac{\dot{X}_t}{X_t} = \beta \ln(\frac{1}{X_t})$. The mixed information source model is $X_t = \frac{1-e^{-\gamma t}}{1+\alpha e^{-\beta t}}$ and corresponds to a diffusion process in which past users as well as a central source diffuse the new technology.

capturing the diffusion of the internet and performs much better than the other two models, which overestimate the rate of diffusion in the late stages.

Accordingly, I use a mixed information source model to model the share of online job posting and write the share of newspaper advertising as

$$s_t^p = 1 - \frac{1 - e^{-\gamma t}}{1 + \alpha e^{-\beta t}} \tag{4}$$

I fit equation (4) to the print share implied by JOLTS data over 2000-2009 and by imposing a share of one in 1994. Figure 4 presents the resulting newspaper share, and Figure 1 plots the new composite HWI along with the original print HWI. Moreover, the new composite HWI does a good job at matching the level of JOLTS job openings over 2000-2009, providing supportive evidence that the MISM can successfully model the share of online advertising.

4 Estimating a matching function

To illustrate this new series, I estimate a matching function over 1951-2009 using Shimer's (2007) measure of the job finding rate as the dependent variable. In a continuous time set-up, the number of new matches m_t at instant t can be modeled by a Cobb-Douglas matching function such that

$$m_t = m_0 u_t^{\sigma_u} v_t^{\sigma_v}$$

with v_t the number of job openings and u_t the number of unemployed. The job finding rate is defined as the ratio of new hires to the stock of unemployed, so that unemployed workers find a job according to a Poisson process with time varying arrival rate $f_t = \frac{m_t}{u_t}$.[8]

I can then estimate the elasticity of the matching function by estimating the following equation

$$\ln f_t = \sigma_v \ln v_t - (1 - \sigma_u) \ln u_t + c + trend + \varepsilon_t. \tag{5}$$

where I construct a measure of the job finding rate as in Shimer (2007) using data on unemployment duration from the Census Population Survey of the BLS.[9]

I estimate (5) with monthly data using the composite HWI from 1951:M01 until 2009:M09. Table 1 presents the result. The elasticity σ_v and σ_u are precisely estimated at respectively 0.32 and 0.49. Interestingly, this result points towards a matching function with decreasing returns to scale. A legitimate concern with this regression exercise is that equation (5) may

[8] A standard assumption at business cycle frequencies is that there are no transitions between participation and non-participation.

[9] More specifically, I recover the job finding probability from $\bar{F}_t = 1 - \frac{u_{t+1} - u_{t+1}^{<1}}{u_t}$ with $u_t^{<1}$ unemployment with duration less than one month. The estimated job finding rate is then given by $f_t = -\ln(1 - F_t)$.

be subject to an endogeneity bias. I then estimate (5) using lagged values of v_t and u_t as instruments.[10] As Table 2 shows, the endogeneity bias is likely to be small as the coefficients are little changed at 0.34 and 0.49. Finally, I follow the rest of the search literature and impose constant returns to scale and obtain $\sigma_v = 0.42$, or $\sigma_u = 0.58$, a value in the middle of the plausible range $\sigma_u \in [0.5, 0.7]$ identified by Petrongolo and Pissarides (2001).[11] This result is very robust across sub-samples, and I obtain essentially the same result over 1985-2009.[12]

5 Conclusion

In this paper, I build a composite Help-Wanted Index that combines job advertising from newspaper and online sources. By modeling the share of online job advertising as the diffusion of a new technology, online job posting and job search, I can combine information on both print and online help-wanted advertising. I use this consistent measure of vacancy posting over 1951-2009 to estimate the elasticity of the matching function and find an estimate of 0.58.

[10] Such instruments are valid if the residual is not serially correlated. The Durbin-Watson statistics for regression (1) in Table 1 is 1.78.

[11] In contrast, Shimer (2005) regresses a variant of (5) after imposing constant returns to scale but obtains a smaller coefficient $\sigma_v = 0.28$. This stems from the fact that Shimer uses the job finding probability, rather the job finding rate, in his regression. However, in a continuous set-up such as that of Mortensen and Pissarides (1994) or Shimer (2005), the equations characterizing the search and matching model involve the job finding rate f_t, not the job finding probability F_t.

[12] Over 1951-1984 (not shown), I get a very similar result with $\sigma_v = 0.41$.

References

[1] Blanchard, O. and P. Diamond. "The Cyclical Behavior of the Gross Flows of U.S. Workers." Brookings Papers on Economic Activity, 1990(2): 85-143, 1990.

[2] Fallick, B. "A Composite Help-Wanted Index," Board of Governors of the Federal Reserve System, Memo, 2008.

[3] Mortensen, D. and C. Pissarides. "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61, pp. 397-415, 1994.

[4] Petrongolo, B. and C. Pissarides. "Looking into the black box: A survey of the matching function," *Journal of Economic Literature*, 39: 390-431, 2001.

[5] Shimer, R. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95(1), pp. 25-49, 2005.

[6] Shimer, R. "Reassessing the Ins and Outs of Unemployment," NBER Working Paper No. 13421, 2007.

Table 1: Estimating a Cobb-Douglas matching function

Dependent variable:	f	f	f	f
Sample	1951:M1-2009:M10	1951:M1-2009:M10	1951:M1-2009:M10	1985:M1-2009:M10
Regression	(1)	(2)	(3)	(4)
Estimation	OLS	GMM	OLS	OLS
σ_v	0.32***	0.34***	0.42***	0.42***
	(0.01)	(0.01)	(0.01)	(0.01)
σ_u	0.49***	0.49***	--	--
	(0.01)	(0.02)		
DW stat	1.78	--	2.10	2.20
R^2	0.87	--	0.85	0.77

Note: All regressions include a linear trend. Standard-errors are reported in parentheses. In equation (2), I use 3 lags of v and u as instruments, and in equation (3) and (4), I allow for first-order serial correlation in the residual.

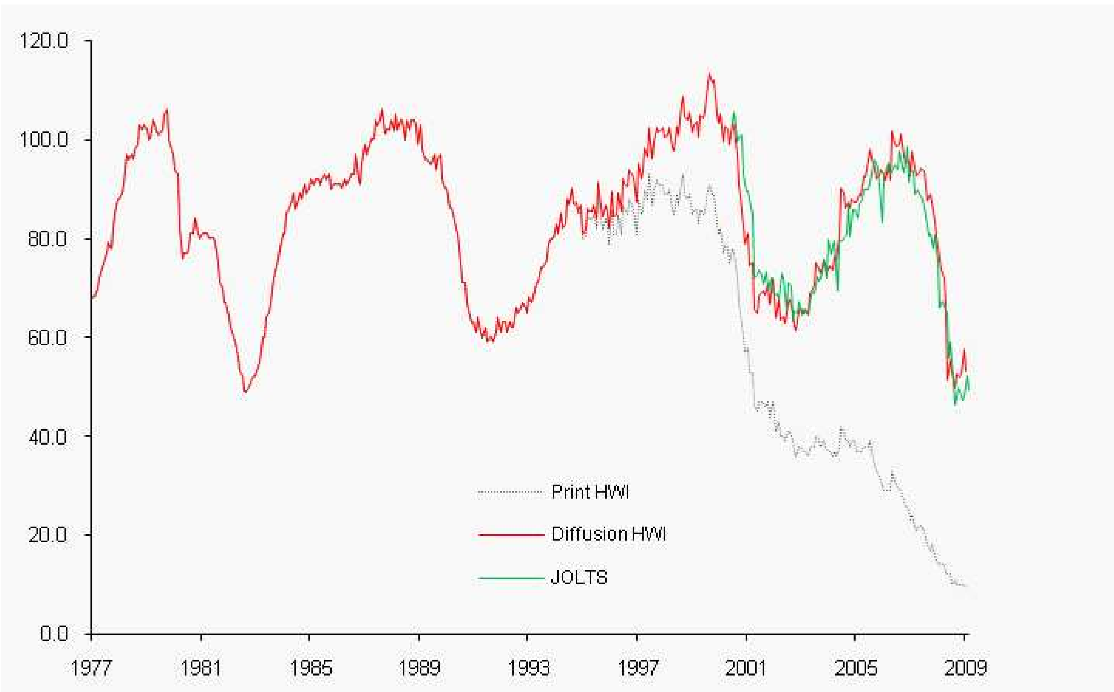

Figure 1: Different indexes of vacancy posting, 1977m01-2009m11.

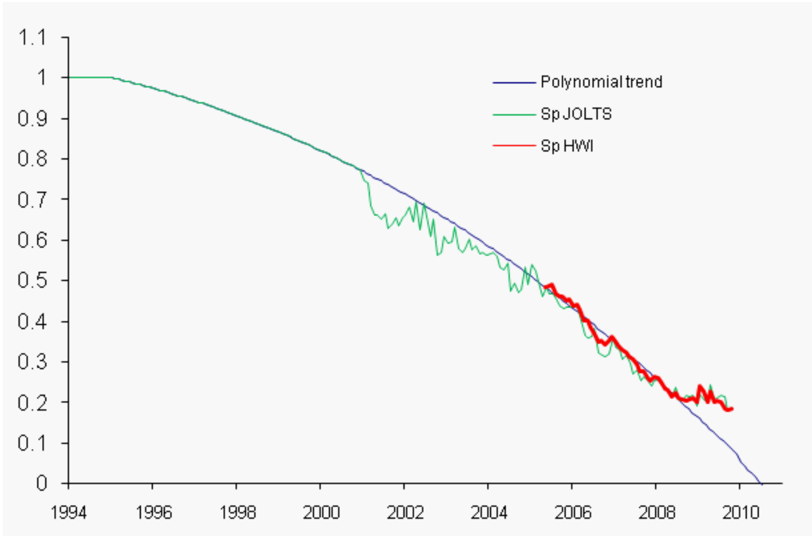

Figure 2: The share of print advertising over 1994-2011, US data and polynomial trend

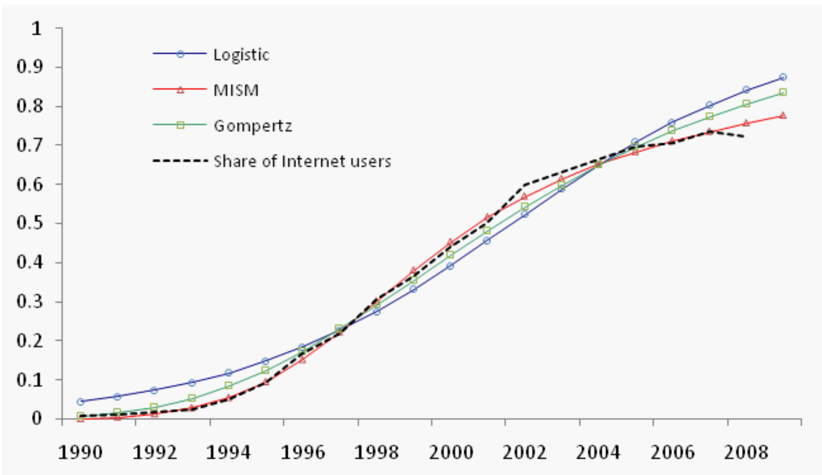

Figure 3: Share of internet users in population, US data and model fits

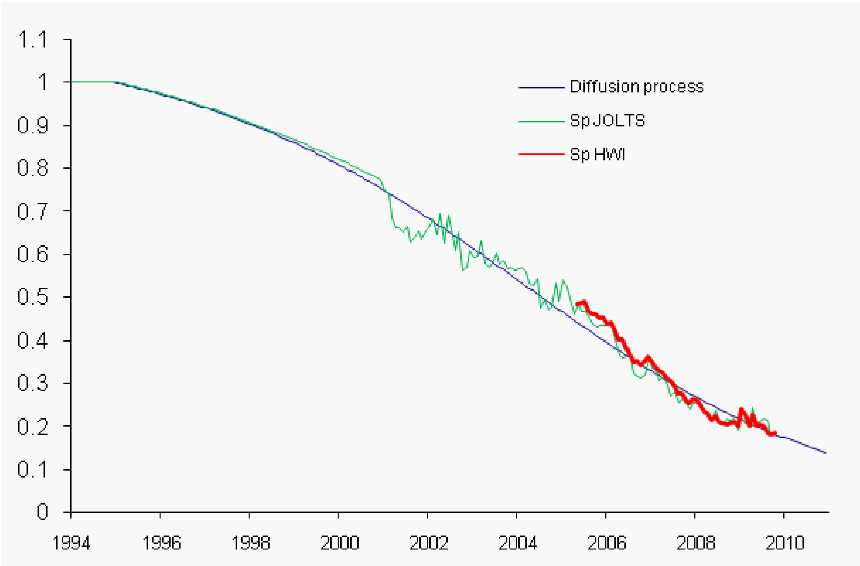

Figure 4: The share of print advertising over 1994-2011, US data and MISM.