Appendix1: Data Smoothing and Stationarity Tests

A Data Smoothing Methods

1 Moving average

To cancel seasonal components from the original series $\{x_t\}$, I take 12-period moving average by setting 2m=12 in the following equation. $\{x_t^*\}$ denotes the adjusted series for $\{x_t\}$.

$$x_t^* = \frac{1}{2m}\left(\frac{1}{2}x_{t-m} + x_{t-m+1} + \ldots + x_{t+m-1} + \frac{1}{2}x_{t+m}\right)$$

In this case, weights are the same except for the first and last observations.

2 Exponential

In case of (single) exponential smoothing, the estimated value for period t is given as follows:

$$x_t^* = (1-\beta)\sum_{j=0}^{t-1}\beta^j x_{t-j}$$

β is a smoothing parameter (0<β<1). In this method, the farther the observations from estimated date t, the less they influence on the estimated variable, and their influence decreases exponentially.

In addition, I tried different versions of exponential smoothing which is defined as follows:

2' Holt-Winters Methods

$$x_t^*(k) = \hat{a}_1(t) + k\hat{a}_2(t)$$
$$\hat{a}_1(t) = (1-\alpha)x_t + \alpha(\hat{a}_1(t-1) + \hat{a}_2(t-1)), 0 < \alpha < 1$$
$$\hat{a}_2(t) = (1-\gamma)(\hat{a}_1(t) - \hat{a}_1(t-1)) + \gamma\hat{a}_2(t-1), 0 < \gamma < 1$$

2'' Seasonal Holt-Winters Methods

Add seasonal component s to the Holt-Winters method as follows:

$$\hat{s}_t = (1-\delta)(x_t - \hat{a}_1(t)) + \delta s_{t-j}, 0 < \delta < 1, j = 12$$
$$x_t^*(k) = \hat{a}_1(t) + k\hat{a}_2(t) + \hat{s}_{t+k-j}$$

3 Seasonal ARIMA process

When monthly data shows systematic seasonal movements, I could model the serial correlation in the disturbance by decomposing it into 3 parts; 1 seasonal autoregressive, 2 seasonal moving average, 3 integration order term. Therefore, I could formulate the series x_t as follows:

$$\phi(L)(1-L)^d(1-L^S)^D x_t = \mu' + \theta(L)u_t$$

The smoothed series show almost the similar characteristics in their movements especially in terms of the existence of unit roots. After we have checked whether the seasonality of series have been removed properly, we select exponentially-smoothed series for the estimation of matching function with regards to M, U, and V. With regards to the market structure index H, we would use the ARIMA process since other smoothing methods do not provide appropriate results because of the limitation in data frequency.

B Panel Unit Root Tests and Cointegration Tests

In this section I discuss the non-stationary panel data tests for unit roots and cointegration. As I showed in section 2, the matching function for region j and period t I would like to estimate is formulated as follows:

$$M_{jt} = A_{jt}[s(H_{jt})]^{\alpha_3} V_{jt}^{\alpha_2} U_{jt}^{\alpha_1}$$
$$A_{jt} = A\exp(\beta_j + \delta_j t + \varepsilon_{jt})$$

A_{jt} is the specification for matching technology and is the element which introduces a stochastic element into model. Vacancy concentration and technology in this model both act as shift factors. Technology includes both a possible intercept β and trend term δ. This model allows for varying intercepts, trends, and varying slopes. Therefore, I could analyze the regional difference in matching speed in the market.

If the data sets are non-stationary, we assume that there exits a long-run relationship between matching, vacancy, and unemployed (and possibly vacancy market structure). An intuitive interpretation of this hypothesis would be that if a long-run relationship between these four variables exists, then including vacancy market concentration index in the matching function specification is reasonable and helpful in describing the growth in matching speed in the long run. Unfortunately, cointegration is a way of analysis when several series are fluctuating in line through a long period, thus it is hard to use this method when we have only short time-series data, such as six years. Therefore, we would try to supplement data shortage and increase the degree of freedom by employing cross-sectional (regional-level) data as well.

As the first step, we need to test whether the series involved contain unit roots or not. If all variables are stationary, then traditional estimation methods can be used to estimate the relationship between variables. If, however, at least one of the series is determined to be nonstationary, then more discussion would be required.

The test statistic we use for stationarity was presented by Im et al. (1997). This statistic tests the null hypothesis of non-stationarity for a variable observed in a panel. The basic idea of this test is that it executes the augmented Dickey-Fuller (ADF) test for each cross-sectional observation, and then assuming each individual cross-sections are independent, combining the result by using large-sample distribution of t-statistics to test the null hypothesis on the panel as a whole. The details are provided as follows.

ADF test in time series can be written as:

$$\Delta y_t = \alpha + \delta t + \rho y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} + v_t$$

where p denotes the number of lags. The null and alternative hypotheses would be written: H_0: yt is an I(1) process versus H_1: yt is an I(0) process. This null hypothesis could be tested using a t-statistic on ρ. Assuming the cross-sections are independent, the average of individual ADF t-statistics would have the following property (Im et al. (1997));

$$\Psi_{\bar{t}} = \frac{\sqrt{N}(\bar{t}_{N,T} - E[\bar{t}_{N,T}(p,0)])}{\sqrt{Var(\bar{t}_{N,T})}} \xrightarrow{converges} N(0,1)$$

$$\bar{t}_{N,T} = (1/N)\sum_{j=1}^{N} t_j$$

tj is the t-statistic for the OLS estimate of ρ for the cross-section of region j, and $E[t_{N,T}(p,0)]$ is taken under null hypothesis ρj=0 for all j and with the choice $p=(p_1, p_2, \ldots, p_N)'$ for the lag-length vector.

The test results are summarized in Table 1-3 . The results are almost the same when we assume the existence of linear individual trends.

Table1: Panel Unit Root Test Result (Im, Pesaran and Shin W-stat)
Null: No unit-root, assuming no linear trends

Variable	Level		First-difference	
Matching	12.68	(1.000)	-36.40	(0.000)
Vacancy	-4.56	(0.000)	-	-
Unemployment	3.11	(0.999)	-7.24	(0.000)

Note: p values are in parentheses. Data series are smoothed by exponential.

Table 2: Panel Unit Root Test Result (Im, Pesaran and Shin W-stat)
Null: No unit-root, without linear trends

Variable	Level1		Level2		Level3		Level4	
Matching	12.58	(1.000)	18.98	(1.000)	1.80	(0.964)	-1.15	(0.125)
Vacancy	-4.38	(0.000)	-3.97	(0.000)	-3.56	(0.000)	-10.55	(0.000)
Unemployment	0.66	(0.744)	-2.34	(0.010)	8.77	(1.000)	0.74	(0.772)
Vacancy share	-		-		-2.65	(0.004)	-5.58	(0.000)

Note: Series are smoothed by HW (1), SHW(2), Arima(3), Moving average(4).

Table3 : Panel Unit Root Test Result (Im, Pesaran and Shin W-stat)
Null: No unit-root, assuming individual linear trends

Variable	Level		First-difference	
Matching	-0.35	(0.363)	-32.37	(0.000)
Vacancy	-1.31	(0.096)	-	-
Unemployment	2.34	(0.99)	-5.73	(0.000)

Note: p values are in parentheses. Data series are smoothed by exponential.

The series of unemployment and matching cannot reject the null-hypothesis of stationarity in the test without a time trend. With regards to vacancy and vacancy share, we cannot reject null hypotheses that vjt (or sjt) for all region j are stationary. On the contrary, unemployment and matching seem to be non-stationary even when we assume linear time-trend for these variables. Given the presence of non-stationary variables in both specifications, we now need to proceed to test for cointegration.

Since vacancy and vacancy share are both stationary, it is not possible that they are cointegrated with unemployment (regressors are not cointegrated with each other).

To test whether the variables of unemployment and matchings could be cointegrated, I employed Nyblom-Harvey (NH) statistics. This statistics test the null hypothesis of 0 common trends among the panel data series against the hypothesis of common trends among variables. Nyblom and Harvey (2000) state the rejection of the null hypothesis of zero common trends implies cointegration.

With regards to unemployment, NH statistics = 24.08, much greater than the one-sided critical value = 10.92 (1 percent critical value), so that the null hypothesis is rejected. At the same time, NH statistics for matching is 23.60, which also could be rejected at the level of 1 percent critical value. Therefore, the estimated coefficient for unemployment on matchings can be interpreted as the potential long-run impacts.