Dynamic Allocation and Pricing in Incomplete Markets: A Survey

Makoto SAITO

Discussion Paper No. 98-E-12
NOTE: IMES Discussion Paper Series is circulated in order to stimulate discussion and comments. Views expressed in Discussion Paper Series are those of authors and do not necessarily reflect those of the Bank of Japan or the Institute for Monetary and Economic Studies.
Dynamic Allocation and Pricing in Incomplete Markets:
A Survey

Makoto SAITO

October, 1998

ABSTRACT. This paper surveys the recent development of empirical and theoretical researches on incomplete markets, thereby pointing out the following respects. First, the theoretical study in this field is motivated by empirical findings of both asset pricing anomalies and heterogenous behavior among economic agents. Second, incomplete insurance combined with either borrowing constraints or transaction costs offers predictions consistent with empirical findings. In addition, the failure of insuring persistent or permanent shocks alone yields empirically reasonable predictions. Third, recent theoretical research has made attempts to endogenize incomplete insurance from first principles. Fourth, incomplete markets may make aggregate shocks distributed disproportionately among agents, thereby having a significant impact on dynamic allocation and pricing. Finally, the theoretical research into incomplete markets triggers a reassessment of welfare implications as to business cycles, economic growth, and financial integration.

JEL classification: G12, D52.
Keywords: incomplete markets, asset pricing.

* Faculty of Economics, Osaka University.
† Correspondence to: Makoto Saito, Faculty of Economics, Osaka University, 1-7 Machikane-yama, Toyonaka, Japan, 560-0043, e-mail: makoto@econ.osaka-u.ac.jp, phone: 81-6-850-5264, fax: 81-6-850-5274.
‡ The author acknowledges an anonymous referee and Yuichi Fukuta for helpful comment. He is also grateful to the Bank of Japan for financial support.
1. Introduction  One of the most important implications of complete markets is that idiosyncratic shocks are shared efficiently among agents. Idiosyncratic shocks may be specific to persons, families, workplaces, industries, or regions in a domestic economy, and particular as to countries in international contexts. Sharing idiosyncratic shocks among heterogeneous consumers implies two immediate theoretical consequences. First, the consumption behavior becomes similar among consumers. In particular, the individual consumption growth is close to one another. Second, consumers demand a moderate premium in exchange for holding risky assets because they are able to diversify both idiosyncratic and aggregate shocks in complete markets.

Empirical research, however, shows that these two predictions are violated profoundly by observations. On the one hand, the asset pricing literature frequently finds that risk premia observed in financial markets are far larger than ones predicted by the complete market framework. On the other hand, panel data studies often observe a substantial heterogeneity in the consumption behavior among individuals. The macroeconomic and microeconomic evidence all suggests that idiosyncratic shocks are not shared efficiently in markets, and that individuals fail to diversify risks. Recent research into incomplete markets is motivated by the above observation of both asset pricing anomalies and heterogeneous consumption behavior.

This paper surveys empirical and theoretical research on incomplete markets with the following emphasis. First, the paper briefly reviews empirical methods to test market completeness within the representative agent framework. In addition, it surveys the recent empirical research which tests predictions specific to incomplete markets using macroeconomic and microeconomic data, in particular, explores how the Euler equation deviates in the absence of completeness.

Second, the paper examines two channels through which incompleteness does matter
Dynamic Allocation and Pricing in Incomplete Markets: A Survey

in affecting dynamic allocation and pricing. One class of theoretical models considers incomplete insurance or the failure of insuring idiosyncratic shocks in markets. In this class of models, main topics include how the nature of idiosyncratic shocks affects asset pricing and how financial frictions such as borrowing constraints and transaction costs limit the ability of self-insurance. The other class takes into consideration how incompleteness makes aggregate shocks distributed disproportionately among different agents. The concentration of aggregate shocks on a subset of the entire population often yields predictions which are fairly different from the complete markets framework. In addition, at more fundamental levels, recent theoretical research has made attempts to endogenize incomplete markets from first principles.

Finally, the paper studies how the welfare assessment of macroeconomic performance is influenced in the absence of completeness. In particular, two topics are discussed: costs of business cycles and welfare implications of economic growth. The welfare evaluation as to both business cycles and economic growth quite often differs substantially between the economy with complete markets and that without ones. Another topic is how welfare distribution changes when an economy moves from less complete markets to more complete.

This paper is organized as follows. Section 2 discusses empirical research, and section 3 deals with theoretical research. Section 4 examines welfare implications of incomplete markets. Section 5 concludes.

2. Empirical research When markets are complete, the intertemporal marginal rate of substitution (hereafter, IMRS) is equalized among agents. The aggregation theorem exploits this equality of the IMRS among consumers, and allows one to construct a representative agent from heterogeneous consumers. Therefore, testing the representative agent framework using aggregate data serves as a starting point for examining market completeness.
2.1. **Aggregation theorem**  
First of all, this subsection briefly reviews the aggregation theorem. Suppose that a market exists for every possible state, and that a contingent claim for state \( \omega(t) \) at time \( t \) is traded at the price \( p(\omega(t)) \) when an economy starts at time 0. Agent \( i \) maximizes the following life-time expected utility at time 0:

\[
\sum_{t=0}^{\infty} \frac{1}{(1+\rho)^t} \sum_{\omega(t) \in \Omega(t)} (\pi(\omega(t))u(c_i(\omega(t))))
\]

subject to

\[
\sum_{t=0}^{\infty} \sum_{\omega(t) \in \Omega(t)} p(\omega(t))(c_i(\omega(t)) + d_i(\omega(t)) - c_i(\omega(t))) = 0,
\]

where \( \rho \) is the rate of time preference, \( \Omega(t) \) denotes a set of all states at time \( t \), \( \pi(\omega(t)) \) is the objective unconditional probability that state \( \omega(t) \) takes place at time \( t \), and \( u \) is the time-additive and state-independent period utility. \( u'(c_i(\omega(t))) > 0 \) and \( u''(c_i(\omega(t))) < 0 \) are assumed. \( c_i, e_i, \) and \( d_i \) indexed by state \( \omega(t) \) imply agent \( i \)'s consumption, exogenous income, and transfer from other consumers at state \( \omega(t) \) respectively.

Assuming that \( \lambda_i \) is the Lagrange multiplier for agent \( i \)'s life-time budget constraint, and deriving the first order condition with respect to agent \( i \)'s \( \omega(t) \)-consumption, we obtain

\[
\frac{1}{(1+\rho)^t} \pi(\omega(t))u'(c_i(\omega(t))) = \lambda_i p(\omega(t)). \tag{1}
\]

The above first order condition (1) with the concavity of the utility function leads to the result that individual consumption increases monotonically in aggregate consumption, or \( c_i(\omega(t)) = f_i(C(\omega(t))) \) with \( f_i'(C(\omega(t))) > 0 \) where \( C(\omega(t)) = \sum_i c_i(\omega(t)) \) (see Wilson [122] and Huang and Litzenberger [67]). Restating this result, individual consumption is influenced by only aggregate shocks which appear at the macroeconomic level. This analytical consequence sheds light on the ability of complete markets to perfectly insure
idiosyncratic shocks.

It is possible to derive from equation (1) the equality of the marginal rate of substitution (MRS) between two different states among consumers. That is, taking two states \( \omega(t) \) and \( \omega(t + 1) \) for example, we obtain, after cancelling out the Lagrange multipliers \( \lambda_i \) and \( \lambda_j \),

\[
\frac{1}{1 + \rho} \frac{u'(c_i(\omega(t + 1)))}{u'(c_i(\omega(t)))} = \frac{1}{1 + \rho} \frac{u'(c_j(\omega(t + 1)))}{u'(c_j(\omega(t)))} \quad \forall \ i \ and \ j.
\]

(2)

In addition, as Rubinstein [100] shows, when the utility function \( u \) belongs to the hyperbolic absolute risk aversion family (the HARA family), the MRS evaluated at individual consumption is equal to that evaluated at aggregate consumption, or

\[
\frac{1}{1 + \rho} \frac{u'(c_i(\omega(t + 1)))}{u'(c_i(\omega(t)))} = \frac{1}{1 + \rho} \frac{u'(C(\omega(t + 1)))}{u'(C(\omega(t)))} \quad \forall \ i,
\]

(3)

where \( C(\omega(t)) = \sum_i c_i(\omega(t)) \). The above equation (3) lays the foundation for a representative agent model as a reasonable approximation to the complete market economy with multiple agents. As shown below, the empirical examination of market completeness is carried out by testing either empirical specifications based on equation (2) using microeconomic data or specifications derived from equation (3) using macroeconomic data.

### 2.2. Tests of capital asset pricing models

The construction of a representative agent relies on the assumption of market completeness; therefore, testing implications available from the representative agent framework can serve as a statistical investigation as to whether markets are complete. The most typical way is to estimate the following form of Euler equation defined on aggregate consumption \( C(t) \):

\[
\frac{1}{1 + \rho} E_t \left( R(t, t + 1) \frac{u'(C(t + 1))}{u'(C(t))} \right) = 1,
\]

(4)
where \( E_t \) denotes the expectation operator conditional on information available as of time \( t \), and \( R(t, t+1) \) implies the holding real gross return on a financial asset between time \( t \) and \( t+1 \). Rubinstein [100] shows that it is possible to rigorously derive equation (4) from equation (3).

The existing literature often assumes the constant relative risk aversion preference (hereafter, the CRRA preference) for a utility function of a representative agent as a reference case, while heterogeneous agents can be replaced by a representative agent under a wide class of time-additive utility functions\(^1\). That is, \( u(C) = \frac{C^{1-\gamma}}{1-\gamma} \), where \( \gamma \) is the degree of relative risk aversion. Hence, estimating equation (4) corresponds to testing a joint hypothesis of a particular functional form of utility as well as market completeness. In this context, market incompleteness is not the whole reason for the rejection of equation (4), but a likely candidate for the interpretation of the rejection\(^2\). We will later explore empirical implications more specific to market incompleteness.

**GMM estimation** A typical estimation procedure for the above Euler equation (4) is the generalized method of moments (hereafter, GMM) proposed by Hansen [49].

According to the GMM estimation result using US financial data\(^3\), two aspects are inconsistent with theoretical implications. First, estimated parameters are frequently out of a reasonable range; the estimated time preference is often negative, while the degree of relative risk aversion is close to zero or negative. Second, while the forecast error of the estimated Euler

\(^1\)Using demand aggregation, Rubinstein [100] shows that if a utility function belongs to the HARA family, then a representative agent can be constructed. Constantinides [27] proves that the capital asset pricing model can be expressed in terms of only aggregate variables under time-additive utility functions with homogeneous beliefs. The assumption about utility functions adopted by Constantinides is weaker than that adopted by Rubinstein in aggregating demand.

\(^2\)Careful readings of the literature indicate that the estimation of the Euler equation using macroeconomic data was not motivated initially by a desire to test market completeness. The main motivation was to examine whether the risk aversion behavior is responsible for the rejection of the efficient market hypothesis which assumes risk neutrality. See Leroy [79].

\(^3\)The empirical study by Hansen and Singleton [52] is a seminal work in this field. Singleton [107] makes an excellent survey on the GMM estimation of the Euler equation.
equation should not include any predictable components under the null-hypothesis, the failure of overidentifying restrictions indicates that forecast errors can be predicted by the current information set. These two findings imply that the representative agent framework with the CRRA preference fails to account for US macroeconomic data.

**calibration methods** Empirical results based on calibration methods are also unfavorable to the prediction of the representative agent framework. A main idea of calibration methods is to simulate a fully-specified general equilibrium version of a representative agent model, and to compare simulated moments with observed moments. Specifying a set of moments under consideration, this method is helpful for checking in which dimension the representative agent framework does not work very well.

The study by Mehra and Prescott [89] is a first example of testing the representative agent model by a calibration method. Their simulation model is an application of Lucas [81]. Adopting the CRRA preference, they specify as a target the unconditional mean of both the equity premium and the risk-free rate. According to their investigation of US financial data, the observed average of the equity premium is far larger than the average predicted by their representative agent model, while the observed average of a risk-free rate is much smaller than the theoretical prediction. Mehra and Prescott themselves call the former inconsistency the equity premium puzzle, and Weil [120] calls the latter the risk-free rate puzzle.

One main source of these puzzles is that the aggregate consumption growth observed in developed countries including US is fairly smooth, and its variance is too small to yield empirically reasonable magnitudes of equity premia and risk-free rates. Mehra and Prescott [89] set up a clear criterion for the appropriateness of a certain class of asset pricing model; a reasonable model should help to resolve the above puzzles. The subsequent literature
also offers a statistical metric between simulated moments and observed ones.\(^4\)

**Hansen-Jagannathan method**  The model evaluation method proposed by Hansen and Jagannathan \([51]\) is similar to the calibration method in spirit. But, the above method, without fully specifying a general equilibrium model, offers a test based on much weaker restrictions than does the calibration method; therefore, the violation of the Hansen-Jagannathan restriction means a much stronger rejection of the model under consideration.

The main idea of the Hansen-Jagannathan method is summarized as follows. Suppose that \(m(t, t + 1)\) denotes the IMRS which discounts one period ahead financial payoffs \((\hat{R}(t, t + 1)\) which is a vector of financial returns) at time \(t\). \(m(t, t + 1)\) should be constructed such that the following unconditional expectation version of Euler equation may hold:

\[
\frac{1}{1 + \rho} E \left( \hat{R}(t, t + 1)m(t, t + 1) \right) = \hat{1},
\]

where \(E\) implies the unconditional expectation operator, and \(\hat{1}\) is a unit vector. In equation (4), therefore, \(m(t, t + 1)\) corresponds to \(\frac{\hat{u}'(C(t + 1))}{\hat{u}'(C(t))}\). A vital part of this method is to assume that \(m(t, t + 1)\) can be approximated by a linear combination of observed \(\hat{R}(t, t + 1)\)'s as

\[
m(t, t + 1) = \text{constant} + \hat{R}(t, t + 1)'\alpha + \text{error terms},
\]

where \(\alpha\)’s are coefficients of a linear regression.

Since \(m(t, t + 1)\) is not directly observable by nature, it is impossible to estimate \(\alpha\)’s in order to exactly recover a systematic part of \(m(t, t + 1)\). It is, however, possible to impose some restrictions on the unconditional moments of \(m(t, t + 1)\) such that equations (5) and (6) may be satisfied at the same time. Hansen and Jagannathan derive the lower bound

\(^4\)See Duffie and Singleton \([36]\). Heaton \([62]\) applies the simulated method of moments to the estimation of non-time-separable preference.
for $\text{Var}(m(t,t+1))$, which corresponds to the variance of a systematic part of equation (6), given a certain level of $E(m(t,t+1))$.

Using US macroeconomic data, Hansen and Jagannathan [51], Cochrane and Hansen [25], and others find that the mean and variance of the IMRS based on the CRRA preference are far outside the region implied by the Hansen-Jagannathan method. These results suggest a much stronger rejection of the representative agent model with the CRRA preference than does the Mehra and Prescott type calibration method$^5$.

**extension of representative agent models** As mentioned before, the failure of representative agent models is caused by either the mis-specification of utility functional forms or the absence of complete markets. Before pursuing as a major topic the latter possibility in the rest of this section, we briefly discuss the former possibility.

Keeping the complete market assumption intact, the literature extends representative agent models in three directions$^6$. First, time-non-separability is introduced into the preference formulation. Constantinides [28] shows that taking into consideration habit formation yields a reasonable magnitude of equity premia using a continuous-time asset pricing model. Abel [1] and Campbell and Cochrane [20] also introduce consumption externality, thereby yielding predictions similar to the habit formation model. Heaton [60], Ferson and Constantinides [40], and others statistically test whether asset pricing models are improved by considering habit and durability using US aggregate data.

Second, the notion of non-expected utility is exploited. Epstein and Zin [38] adopt non-expected utility in defining asset pricing models. Epstein and Zin [39] estimate the Euler equation based on their non-expected utility model using US financial data; their estimation

---

$^5$The subsequent development of the Hansen-Jagannathan method considers statistical inferences (Hansen, Heaton, and Luttmer [50]).

$^6$Kocherlakota [75] systematically surveys attempts to resolve the equity premium puzzle and the risk-free rate puzzle. Campbell, Lo, and MacKinlay [21] offer a comprehensive assessment of the recent econometric development in this field.
result is a little mixed and quite sensitive to the choice of instrumental variables. Third, catastrophic events such as the Great Depression are treated explicitly in specifying a stochastic environment. Using calibration methods, Reitz [99], Cecchetti, Lam, and Mark [22], and others find that large equity premia are predicted by asset pricing models with catastrophic events.

2.3. Tests of insurance using microeconomic data Testing microeconomic implications using equation (2) is another way to examine market completeness, and offers a more direct evidence for whether markets are complete. As first pointed out by Townsend [114], a striking feature of equation (2) is that it holds for any realized state as long as contingent claims are available for all possible states. Under the null hypothesis of complete markets, hence, equation (2) impose rather tight restrictions on the realized path of individual consumption.

More concretely, the realized consumption growth is exactly equal among consumers under the CRRA utility \((u(c) = \frac{1-e^{-\gamma}}{1-\gamma})\). That is,

\[
\ln c_i(w(t+1)) - \ln c_i(\omega(t)) = \ln c_j(\omega(t+1)) - \ln c_j(\omega(t)) \quad \forall \ i \text{ and } j.
\]  

(7)

Similarly, the first difference in consumption is equalized among consumers under the constant absolute risk aversion preference \((u(c) = -\frac{1}{\sigma} \exp(-\sigma c), \text{hereafter, the CARA preference})\):

\[
c_i(w(t+1)) - c_i(\omega(t)) = c_j(\omega(t+1)) - c_j(\omega(t)) \quad \forall \ i \text{ and } j.
\]  

(8)

Given more desirable features of the CRRA utility in the uncertainty context (see Huang and Litzenberger [67]), the CRRA preference is often preferred to the CARA preference in empirical practice. Following such practice, this paper mostly focuses on the case with the CRRA preference.
Exploiting the aggregation theorem, Mace [85] and others derive from equation (3) with the CRRA preference the following alternative specification for testing full insurance:

\[ \ln c_i(w(t + 1)) - \ln c_i(w(t)) = \ln C(\omega(t + 1)) - \ln C(\omega(t)) \quad \forall i. \]  

(9)

The above equation throws a clear light on the implication that the individual consumption growth is determined by only a common factor, or the aggregate consumption growth.

Examining empirical implications such as equations (7), (8), and (9) provides a simple, but powerful tool for a test of complete markets or full insurance. Specifically, such an examination tests whether individual consumption growth is influenced by a common time-specific factor (the aggregate consumption growth) or by person-specific factors. In transforming these theoretical implications into statistical relationships, researchers usually introduce either measurement errors or individual preference shocks as statistical error terms.

**panel data study** While many empirical papers test full insurance implicitly as pointed out by Hayashi [56], more and more papers have conducted an explicit test of full insurance using US panel data. The examples using the Panel Study Income Dynamics (PSID) include Altug and Miller [7], Cochrane [24], Altonji, Hayashi, and Kotlikoff [6], McCarthy [88], and Hayashi, Altonji, and Kotlikoff [58], while those using the Consumer Expenditure Survey (CES) Mace [85] and Nelson [91]. In addition, Townsend [115], [116], and others examine household panel data from developing countries.

Overall, empirical results reject full insurance. Careful readings of these results illuminate how full insurance is rejected. While Mace [85] is in favor of full insurance, Nelson [91] discusses that the result of Mace is sensitive to a more accurate measurement of consumption and employment. Cochrane [24] shows that individual consumption grows more
slowly in response to long illness and involuntary job loss.

McCarthy [88] demonstrates that the relevance of full insurance differs between high-
While Altug and Miller [7] argue that ignoring nonseparability of utility in consumption
and leisure is responsible for the rejection of full insurance under separability, Hayashi et al.
[58] construct a test such that it has a statistical power against self-insurance as alternative,
thereby rejecting full insurance even under nonseparability.

differentiation between full insurance and self-insurance Even in the absence
of a full set of contingent claims, consumers are still able to self-insure income fluctuation
to some extent by saving and dissaving in financial and physical assets. Under the self-
insurance hypothesis, the following Euler equation still holds at the individual level:

\[
\frac{1}{1 + \rho} E_t \left( R(t, t+1) \frac{u'(c_{i(t+1)})}{u'(c_{i(t)})} \right) = 1 \quad \forall \ i.
\]  

(10)

One thing to be noticed immediately is that equation (10) is a much weaker restriction
than is equation (2). Accepting equation (10) cannot rule out rejecting equation (2) at
all. A typical way to statistically test equation (10) using panel data is to predict a
component of idiosyncratic shocks to individual resources which are realized between time
t and t + 1 using the time t information set as an instrument, and to examine whether the
one-period ahead forecast error of the Euler equation is correlated with such a predictable
(instrumented) component\(^7\). More concretely, taking for example a change in individual
wages between time t and t + 1 (\(\Delta w_i(t + 1)\)), the above test consists of first instrumenting
\(\Delta w_i(t + 1)\) by \(\Delta w_i(t)\) and other current variables, and second testing whether the realized

\(^7\) Hall [46] is a seminal paper which first proposed such a test for the self-insurance, although his empirical
exercise did not use microeconomic data, but macroeconomic data.
forecast error between time $t$ and time $t+1$ is correlated with this anticipated component of $\Delta w_i(t+1)$.

Given the above empirical procedure, the absence of the correlation between the forecast error of the Euler equation and the idiosyncratic resource shock anticipated by the current information set supports the self-insurance hypothesis, but not necessarily the full insurance hypothesis. Putting emphasis on this difference, Hayashi, Altonji, and Kodikoff [58] propose to include not only the current information, but also the future information in the set of instrumental variables in order to differentiate full insurance from self-insurance.

In addition, testing the self-insurance using aggregate data calls for due caution. As discussed in the previous section, the failure of complete markets leads to the failure of the Euler equation defined on aggregate consumption (equation (4)). Hence, incomplete markets or the failure of full insurance is a likely candidate for the interpretation of the rejection of equation (4) by aggregate data.

The rejection of equation (4), however, does not necessarily imply that of equation (10). There is still a room for equation (10) to be satisfied at the individual level as far as the self-insurance hypothesis holds. Let us follow Attanasio and Weber [13] to illuminate this point. As Hall [47] demonstrates, assuming the CRRA utility and linearly expanding equation (4) up to a low order, we obtain

$$\ln C(t + 1) - \ln C(t) = \text{constant} + \frac{1}{\gamma}(R(t, t + 1) - \rho) + \text{error term.} \quad (11)$$

The same expression is available from equation (10) for the individual level:

$$\ln c_i(t + 1) - \ln c_i(t) = \text{constant} + \frac{1}{\gamma}(R(t, t + 1) - \rho) + \text{error term} \quad \forall \ i.$$
Summing the above equation over all consumers leads to

\[
\frac{1}{I} \sum_{i} (\ln c_i(t + 1) - \ln c_i(t)) = \text{constant} + \frac{1}{\gamma} (R(t, t + 1) - \rho) + \text{error term},
\]

(12)

where \( I \) denotes the number of consumers. Notice that the error term of equation (12) includes only aggregate shocks because idiosyncratic shocks are averaged out.

When full insurance holds, we obtain \( \frac{1}{I} \sum_{i} (\ln c_i(t + 1) - \ln c_i(t)) = \ln C(t + 1) - \ln C(t) \) from equation (3), and the above two Euler equations (11) and (12) hold at the same time. When full insurance fails, but self-insurance holds, however, \( \frac{1}{I} \sum_{i} (\ln c_i(t + 1) - \ln c_i(t)) \) is no longer equal to \( \ln C(t + 1) - \ln C(t) \), and only the latter Euler equation (12) holds.

Attanasio and Weber [13] argues that, when the self-insurance hypothesis is tested by the aggregate-level-specification, the left hand side of the former Euler equation (\( \ln C(t + 1) - \ln C(t) \)) should be replaced by \( \frac{1}{I} \sum_{i} (\ln c_i(t + 1) - \ln c_i(t)) \) in order to consider the consumption heterogeneity due to market incompleteness.

**liquidity constraints** Liquidity constraints keep consumers from financing current consumption, and limits the ability of consumers to self-insure income fluctuation. In this sense, it is crucially important for the possibility of liquidity constraints to be taken into consideration. The literature on liquidity constraints is, on the other hand, extremely large, therefore, in what follows we will discuss only issues viewed as relevant from the perspective of this paper. Hayashi [56], Deaton [31], and Attanasio [10] offer a comprehensive survey on theoretical and empirical studies of liquidity constraints.

The most important empirical implication of liquidity constraints is that the Euler equation (10) fails to hold at the individual level. More specifically, the individual consumption growth is steeper than that implied by the optimal allocation, and the left hand side of the
Euler equation is smaller than one (e.g. see Deaton [30]):

\[
\frac{1}{1 + \rho} E_t \left( R(t, t + 1) \frac{u'(c_{t+1})}{u'(c_t)} \right) \leq 1.
\]

Zeldes [123] empirically explores the above implication as to the deviation of the Euler equation using microeconomic panel data. He finds that the Euler equation does not hold for a fraction of the population, in particular for low-wealth households. Runkle [101], however, presents the empirical results which are quite different from those of Zeldes [123].

The evidence for liquidity constraints, or the empirical rejection of the Euler equation at the individual level, gives rise to somewhat acute tension in regard to which model to consider first in empirical practice. If not only full insurance, but also self-insurance fails to hold, researchers are forced to give up tight theoretical restrictions such as equations (2) and (10) in specifying empirical relationships. There may be three responses to the possibility of liquidity constraints.

First, one may explicitly solve for an equilibrium of the economy with neither full insurance nor self-insurance. Since there is no closed-form solution in most cases, a wide array of numerical simulation techniques are adopted as a means of solving models. Several examples for this approach will be discussed in Section 3.

Second, one may want to consider the case where liquidity constraints are binding (the self-insurance hypothesis fails to hold), but contingent claims are available for all possible states (markets are complete). One twist here is that this approach is still able to exploit the aggregation property due to the assumption of complete markets. Thanks to this, the

---

8 Hayashi [54] explores the possibility of liquidity constraints by estimating a consumption function using US cross-sectional data.

9 Runkle [101] suggests that either Zeldes' usage of family-specific dummies, the statistical inference about the moving-average error structure, or a difference in sample periods is responsible for the difference in the empirical results.
liquidity constraint at the individual level yields some empirical implications for aggregate data. Following the above approach, He and Modest [59] and Luttmer [84] impose the Hansen-Jagannathan [51] type restriction on aggregate data for the case where individual consumers are subject to liquidity constraints\(^\text{10}\).

Third, it may be fair to point out at this moment that some theoretical considerations still make the self-insurance hypothesis the reasonable and powerful approximation of consumption behavior. As Hayashi [56] discusses, when the utility level at zero consumption is substantially low, consumers arrange their consumption plan such that the zero or negative consumption state may be avoided; consequently, liquidity constraints may be never binding in optimal consumption plans. In addition, by simulation, Zeldes [124] presents the case where zero consumption is avoided along the optimal consumption path under the CRRA preference. In what follows, this paper will take this line of theoretical reasoning at some points, and consider the self-insurance hypothesis as one of reasonable assumptions.

**cross-sectional data study** One of the most serious problems empirical researchers often face in testing either full insurance or self-insurance is that it is extremely hard to obtain high quality panel data. In the case of US, for example, the PSID traces the same individual for long periods, but it records only food consumption. This data restriction

\(^{10}\)More concretely, He and Modest [59] assumes the following notion of a liquidity constraint, called a market wealth constraint, in the complete market setup:

\[
\sum_{\tau=t}^{\infty} \sum_{\omega(\tau) \in \Omega(\tau)} \pi(\omega(\tau)) R_t(\omega(\tau)) \geq 0,
\]

where \(R_t(\omega(\tau))\) is the time \(\tau\) payoff in the subsequent time periods of the entire portfolio (including financial, physical, and human assets) constructed at time \(t\). Other notations are the same as ones in the main text. The above constraint implies that the present value of the consumer’s portfolio must be nonnegative. Luttmer [84], on the other hand, uses a stronger notion of a liquidity constraint, called a solvency constraint, as follows:

\[
R_t(\omega(\tau)) \geq 0,
\]

for any state in the future. That is, any portfolio that includes debt in some future states is prohibited.
Dynamic Allocation and Pricing in Incomplete Markets: A Survey

makes it impossible to test equations (2) and (10) using natural definitions of consumption such as non-durable and service expenditures.

To overcome this problem, Deaton and Paxson [32] propose to explore the cross-sectional implication of equation (10) using the consumption inequality measure within the same cohort. Their idea is simple, but appealing. As persistent or permanent uninsured shocks hit upon individual resources, and are accumulated over time, the cross-sectional dispersion of consumption within a fixed cohort should grow with age. When idiosyncratic shocks are pooled in insurance markets, on the contrary, the within-cohort consumption inequality should be flat with respect to age.

Ohkake and Saito [94] demonstrate that the variance of logarithmic consumption serves as the exact within-cohort inequality measure which is theoretically consistent with equation (10) with the CRRA preference. In addition, Deaton and Paxson [32] and Ohkake and Saito [94] argue that the comparison between the within-cohort inequality of consumption and that of disposable income enables one to test a simple version of liquidity constraints where the current consumption is equal to the current disposable income.

Deaton and Paxson [32] indeed find that the consumption inequality grows quickly with age within a fixed cohort using the cross-sectional data of the US, the UK, and Taiwan. Their findings suggest that a large component of idiosyncratic shocks is not shared efficiently in the insurance markets of these countries. Ohkake and Saito [94] add the case of Japan to the list of the evidence for the failure of full insurance.

Lacking in diversification of financial assets Casual observations indicate that not only idiosyncratic shocks, but also aggregate shocks fail to be shared among the entire population in markets. Because investors and consumers can diversify portfolios in complete
markets, such observations present evidence against complete markets\textsuperscript{11}. Several household studies, including Avery, Elliehausen, and Kennickell [14], find that a substantial fraction of households fail to diversify financial assets efficiently and keep holding conservative and safe portfolios, while a fairly small portion of the population hold most shares of risky assets such as corporate equities. These findings suggest that aggregate shocks, both capital gains and losses, concentrate on the limited number of households who hold equities and other risky assets in their portfolio.

The above evidence for the absence of portfolio diversification is dramatically contradictory to the theoretical implication of asset pricing models which underlie the representative agent framework as well as self insurance. In theory, consumers have a strong incentive to earn higher average returns by diversifying aggregate risks. In practice, however, many households miss such investment opportunities for some reasons.

Mankiw and Zeldes [87] demonstrate that the fact that many households do not hold equity shares is responsible for the failure of the representative agent model. Splitting the sample of the PSID between those holding equities and those not holding, they find that the Euler equation holds more consistently in the stockholder sample than in the non-stockholder sample. Attanasio, Banks, and Tanner [11] make the same type of investigation using UK microeconomic data. These estimation results call for theoretical models with disproportionate allocation of aggregate shocks.

2.4. Evidence from international finance When agent $i$ is interchanged with country $i$, all implications of the above framework can be immediately carried over to the context of international finance. The equality of consumption growth among consumers is now equivalent to the consumption correlation among countries, while a lack of diversification of

\textsuperscript{11}One caveat here is that the diversification of financial portfolios does not necessarily imply complete markets. Even if there are no insurance markets for idiosyncratic shocks, investors and consumers could control aggregate risks (market risks) by portfolio diversification.
household financial portfolios now corresponds to a preference for home equity over foreign equity. According to Hayashi et al. [58], indeed, the earliest test of risk-sharing due to Leme [78] was in the context of the cross-country comparison of consumption growth.

Backus, Kehoe, and Kydland [15] document a number of discrepancies between theory and data for domestic and international aspects of business cycles. The most striking discrepancy in regard to international risk sharing is a lack of consumption correlation across countries. According to Obstfeld [93] and van-Wincoop [118], a lack of consumption correlation across countries implies a potential risk-sharing opportunity, and all countries can reap large welfare gains from international risk-sharing\textsuperscript{12}. In other words, arbitrage opportunities are not exploited in international markets. In the literature, the above discrepancy is often called the consumption correlation puzzle.

French and Poterba [41] first document a preference of investors for home equity over foreign equity in stock markets. Tesar and Werner [110] argue that there is a strong home bias in national investment portfolios despite the potential gain from international diversification. In the literature, the above phenomenon is called the home bias puzzle\textsuperscript{13}. The consumption correlation puzzle and the home bias puzzle jointly indicate that there is a potential room for country-specific shocks to be shared more effectively in international stock and insurance markets.

2.5. Studies based on Japanese data To conclude this section, we briefly review empirical studies based on Japanese data. First of all, let us examine tests of capital asset pricing models (CAPM) using aggregate data. Contrary to the result of the case of US, the forecast error of the Euler equation for stock returns does not contain predictable components. Using monthly data, Hamori [48] and Hori [66] pass the over-identification

\textsuperscript{12}In contrast, as discussed in Section 4, Tesar [109] argues that the gain from international risk-sharing is not significantly different from zero.

\textsuperscript{13}Lewis [80] surveys the literature on the home bias puzzle.
test. Their estimation, however, finds risk-neutrality or near risk-neutral behavior, therefore, stochastic discount factors (equivalent to the IMRS's) play a rather limited role in accounting for predictable components of stock returns. Using semi-annual data, on the other hand, Nakano and Saito [90] find that consumers are indeed risk averse (the degree of RRA is about 2.5).

The above results favorable for the CAPM do not necessarily support the representative agent framework. Nakano and Saito [90] show that, allowing for the generalization of stochastic discount factors, any single stochastic discount factor is not able to explain multiple asset returns consistently. In addition, Hori [66] finds that the unconditional moments of stochastic discount factors is outside the Hansen-Jagannathan region, thereby indicating that the representative agent framework is rejected profoundly by Japanese markets data.

There are only a few papers based on Japanese panel data. Hayashi [55] uses the 1982 Survey of Family Consumption compiled by the Economic Planning Agency, in which families are interviewed every three months for one year, and tests the permanent income hypothesis with due consideration for liquidity constraints. He finds that the permanent income hypothesis (the self-insurance hypothesis) applies to a large fraction of the population, and that only 15 percent of the population are liquidity-constrained.

Kohara [76] conducts a test of full insurance using the 1993 and 1994 Japanese Panel Surveys of Consumption (JPSC) compiled by the Institute of Household Economy. Like the PSID, the JPSC surveys the same household once a year. In addition, Ohtake and Saito [95] test full insurance exploiting the panel data structure of the 1995 Family Income and Expenditure Survey conducted by the Statistics Bureau of the Japanese Government. This survey interviews the same household every month for half a year. Both Kohara [76] and Ohtake and Saito [95] reject the full insurance hypothesis strongly.
Dynamic Allocation and Pricing in Incomplete Markets: A Survey

Ohtake and Saito [94] conduct the Deaton and Paxson [32] type examination using the cross-sectional household data called the National Survey on Family Income and Expenditure compiled by the Statistics Bureau of the Japanese Government. As mentioned before, they find that consumption inequality grows quickly with age within a fixed cohort, thereby suggesting that permanent idiosyncratic shocks remain uninsured in the economy.

In Japan, the examination on regional risk-sharing is particularly interesting because the fiscal system is expected to serve as a risk-sharing scheme by a means of transferring resources among prefectures. Using prefectural data, van Wincoop [119] tests whether risks are shared effectively among regions, and finds quite low consumption correlations among prefectures. His finding suggests that prefectures are still subject to uninsured region-specific shocks despite the public regional transfer. In regard to the home bias puzzle, Kang and Stulz [72] examine foreign stock ownership in Japan. They find that foreign investors display some patterns of preferences for Japanese stocks, which are not necessarily consistent with the optimal portfolio choice.

3. Theoretical models In the past decade, theoretical models have been constructed in order to account for the observations discussed in the previous section, that is, asset pricing anomalies and heterogeneous consumption behavior. In particular, the literature has put emphasis on three respects: incomplete insurance, endogenous constraints in dynamic models, and disproportionate allocation of aggregate shocks. This section surveys the recent theoretical development model by model

3.1. Incomplete insurance Let us begin with the case where any claims contingent on person-specific events are not available in markets, while claims contingent on macroeconomic events are traded without any frictions in financial markets. Restating this, insurance

---

14 Heaton and Lucas [64] survey the literature of asset pricing models with market imperfections.
markets themselves are missing, but consumers are allowed to self-insure person-specific shocks by freely having access to financial markets.

While the absence of insurance markets is often justified by either moral hazard, adverse selection, or limited commitment, missing insurance is treated as exogenous constraints for the moment. Issues as to endogenous constraints will be discussed in the next subsection. In what follows, what is implied by ‘an idiosyncratic shock’ is a shock to which the law of large number is applicable at the aggregate level.

As a literature remark, a dynamic model with both incomplete insurance and liquidity constraints was initiated by Bewley [19]. His model considers only idiosyncratic shocks abstracting aggregate shocks, and suggests that a risk-free rate lies below the rate of time preference. Hence, his finding has already suggested that this line of research has a chance to resolve at least the risk-free rate puzzle. More recent theoretical work has made attempts to evaluate the extent to which the risk-free rate puzzle is resolved as well as to analyze the possibility of resolving the risk premium puzzle by introducing aggregate shocks.

**transitory shocks versus permanent shocks** How different the prediction of this class of models is from that of the standard representative agent framework depends mainly on how much consumers can self-insure idiosyncratic shocks. If the self-insurance ability is inadequate, then the theoretical prediction differs substantially between the two models. Otherwise, there may be little difference. The recent literature finds that the self-insurance ability relies on the nature of idiosyncratic shocks on individual endowment.

Telmer [108], Lucas [83], and Heaton and Lucas [63] show that transitory idiosyncratic shocks are self-insured almost perfectly, and that models with incomplete insurance cannot generate any predictions improving on what the representative agent framework predicts. Constantinides and Duffie [29], on the other hand, demonstrate that agents are not able to effectively self-insure permanent idiosyncratic shocks on individual endowment. Saito
[104] presents a similar case using a continuous-time asset pricing model. As discussed below, their models with permanent idiosyncratic shocks differ from the representative agent models in several respects. Constructing a two-period model, Weil [121] also presents the case where incomplete insurance matters in asset pricing; his model yields predictions similar to the above dynamic models with permanent shocks.

The main reason why theoretical predictions depend on whether idiosyncratic shocks on individual endowment is transitory or permanent is summarized as follows. In the case of transitory shocks, in response to negative shocks, consumers can borrow from markets or draw down assets without any drop in consumption because such shocks have negligible impact on the present value of life-time income. Similarly, consumers receiving positive shocks lend to markets instead of raising consumption because of very little effect on human capital. Such borrowing and lending recover an equilibrium in financial markets. Transitory idiosyncratic shocks, therefore, do not have any significant impact on the individual consumption plan, and consumers still behave alike under adequate self-insurance.

In the case of permanent shocks, on the other hand, idiosyncratic shocks have a substantial impact on individual human capital, and individuals change current consumption in response to realized idiosyncratic shocks instead of keeping consumption intact by a means of financial trading\(^\text{15}\); positive (negative) shocks raises (lowers) consumption. Hence, consumers are subject to uninsured shocks, and behave differently. Exposure to uninsurable shocks triggers precautionary saving on the side of individual consumers. This impact on saving brings about immediate effects on asset pricing on the one hand, and on capital accumulation and economic growth on the other. We first review the asset pricing implica-

\(^{15}\) Constantinides and Duffie [29], Saito [104], Weil [121], all of these models yield an equilibrium without any financial trading. Krussell and Smith [77] demonstrate that this kind of the no-trading outcome generates unrealistic distributions of asset holdings. The no-trading outcome may be interpreted as quite analytically tractable and abstracting from other trading motives.
tion of incomplete insurance, while the prediction for capital accumulation and economic growth later in Section 4.

Precautionary savings triggered by permanent shocks generate overall demand for assets, thereby lowering asset returns. For this reason, the Euler equation (4) does not hold at the macroeconomic level (notice that equation (10) still holds at the individual level under the self-insurance hypothesis). As Constantinides and Duffie [29] and Saito [104] show, the left hand side of equation (4) is less than one:

\[
\frac{1}{1 + \rho} E_t \left( R(t, t + 1) \frac{u'(C(t + 1))}{u'(C(t))} \right) < 1.
\]

The above-mentioned two papers parameterize the difference between both sides of the above inequality under the assumption of stochastic processes.

As mentioned before, the above prediction for lower asset returns is suggestive for the risk-free rate puzzle, or why risk-free rates are so low. Saito [104] predicts a reasonable magnitude of a risk-free rate under plausible parameters. This class of models is, however, somewhat mixed in resolving the equity premium puzzle. Grossman and Shiller [45] has already shown that equity premia \( (R(t, t + 1) - R_f(t, t + 1) \) where \( R_f \) is a risk-free rate) are orthogonal to the IMRS on the average at the aggregate level even in the presence of uninsured idiosyncratic shocks, assuming the diffusion process for individual consumption. That is,

\[
\frac{1}{1 + \rho} E_t \left( (R(t, t + 1) - R_f(t, t + 1)) \frac{u'(C(t + 1))}{u'(C(t))} \right) = 0.
\]

The above Euler equation implication is exactly identical to that of the representative agent framework, therefore, there is no impact of incomplete insurance on risk premia in this case. Saito [104] emphasizes the above result.

Weil [121], however, finds that larger equity premia are generated in his model. Con-
stantinides and Duffie [29], allowing for the stochastic volatility for idiosyncratic shocks, present the case where the left hand side of the above Euler equation is not equal to zero. In particular, they find that it is positive, and equity premia are larger in incomplete insurance than in complete markets when the variance of idiosyncratic shocks is counter-cyclical. Their finding is corresponding to the implication from the two-period model of Mankiw [86]. In addition, Saito [105] demonstrates that, if the preference is switched from the CRRA to the non-expected utility (due to Epstein and Zin [38]), then equity premia is not orthogonal to the IMRS on the average even at the aggregate level, thereby showing the effect on equity premia of the interaction between non-expected utility and incomplete insurance.

In summary, the implication for the equity premium puzzle may depend on either underlying stochastic processes of idiosyncratic shocks or preference functional forms. As will be discussed in Section 4, the effect of incomplete insurance on equity premia is also closely related to the evaluation of the cost of business cycles.

**liquidity constraints, short sales constraints, and transaction costs**  The lesson from the previous subsection is that, if idiosyncratic shocks on individual endowment are transitory, then these shocks are effectively smoothed by self-insurance; the asset pricing implication is similar to that predicted in the representative agent framework. This reasoning implies that, once market frictions such as liquidity constraints, short sales constraints, and transaction costs limit the self-insurance ability, there may be a reasonable improvement in prediction even under transitory idiosyncratic shocks.

When market frictions are introduced into asset pricing models, the Euler equation (10) no longer holds at the individual level. Consequently, tight theoretical restrictions are lost immediately, models may not be analytically tractable, and no closed-form solution may be available. For this reason, the research works in this field, including Aiyagari [3], Aiyagari
and Gertler [5], Den Haan [33], Heaton and Lucas [63], [65], Huggett [68], Krusell and Smith [77], Lucas [83], and Telmer [108], adopt sophisticated numerical calculation techniques. Both high quality personal computers and flexible program codes enable researchers to use such sophisticated techniques.

Lucas [83] and Telmer [108] demonstrate that mild borrowing and short sales constraints have little impact on the self-insurance ability, and that consumers are still able to effectively smooth transitory idiosyncratic shocks. According to their further simulation, a reasonable prediction of equity premium requires both strong degrees of these constraints and persistency of idiosyncratic shocks.

In regard to transaction costs, Aiyagari and Gertler [5] show that liquidity premia over returns on costlessly-transacted assets are influenced by bid-ask spreads and the transaction probability; their model obtains non-negligible premia from simulating the transaction probability with observed bid-ask spreads. Heaton and Lucas [65] decompose the effect of transaction costs on equity premia into two components; a direct effect due to transaction margins and an indirect effect due to an increase in individual consumption volatility; the latter effect corresponds to that due to a reduced ability of self-insurance. They find that the direct effect dominates, and that the model predicts a sizable equity premium only if transaction margins are large.

3.2. **Endogenous constraints in dynamic models**

Both the frequent empirical rejection of the full insurance hypothesis and the reasonable theoretical effect brought about by incomplete insurance as exogenous constraints raise a more fundamental question: why are markets incomplete? These two different fundamental reasons may yield similar empirical implications, but suggest different normative implications. Hence, theorizing incomplete insurance from first principles is necessary for writing proper policy recommendations. Insurance markets may break down due to the following fundamental factors: adverse
selection, moral hazard, and limited commitment. This subsection considers the recent development of dynamic models where market incompleteness emerges endogenously because of either of the latter two factors.

**informationally-constrained insurance contract** Consider the case where the individual income is subject to idiosyncratic shocks, but such shocks are not observable (verifiable). In this case, any insurance contract contingent on the report of the insured as to his income may be subject to the problem of moral hazard; the insured has an incentive to under-report his income to receive compensation. As discussed in the previous subsection, individuals are able to insure income fluctuation to some extent by lending and borrowing. Townsend [113], however, points out that there are efficient incentive-compatible schemes which can dominate the self-insurance by a simple loan contract. Green [42], Thomas and Worrall [112], and others explore the possibility of such efficient insurance contracts using the dynamic programming approach.

A vital idea of their model is that a contract is conditioned not only on the current report of income, but also on the history of the past reports. Given this conditioning, an insurance contract can penalize the insured who report low income by lowering their future compensation, and it can promise those reporting high income that they will receive more compensation in future periods. As a result, such a scheme can solve the problem of incentive compatibility. Green and Oh [43] examine the empirical implication for the IMRS of this kind of informationally-constrained efficient contract, and compare it with that of the permanent income hypothesis (the self-insurance hypothesis) and models with liquidity constraints. Atkeson and Lucas [8] discuss the above efficient contract in a general equilibrium setup16. Phelan [97] introduces aggregate shocks into the setup of the

---

16Green [42], Thomas and Worrall [112], and others do not take into consideration macroeconomic resource constraints, consequently, their analysis is a partial equilibrium approach.
informationally-constrained efficient contract.

The analytical result has been so far available for the case of independently and identically distributed idiosyncratic shocks (i.i.d. shocks or purely transitory shocks) with the CARA preference. As discussed in the previous subsection, self-insurance works quite effectively for insuring transitory shocks. Therefore, the improvement of the above efficient contract over self-insurance may be marginal in terms of efficiency gains, and empirical implications for the consumption allocation may not differ substantially between the two. If the efficient insurance scheme allows for persistent or permanent idiosyncratic shocks, the empirical implications of this class of models are expected to be differentiated from those of the self-insurance hypothesis.

**self-enforcing insurance contract** Another reason for the absence of full insurance is limited commitment. In the complete market paradigm, consumers are assumed to commit themselves to any contract forever, once they agree upon it at time 0. It may be, however, next to impossible to enforce initial contracts on consumers perfectly even under a sophisticated legal system; consumers always have an incentive to renege upon initial contracts. Given such limited commitment, contract should be made such that there is no incentive to breach it. In other words, any contract should be self-enforcing. The consumption allocation may differ between with and without self-enforcing constraints.

Hayashi [57] argues that limited commitment is more responsible for the absence of full insurance than moral hazard by pointing out that a fraction of aggregate consumption is often correlated with publicly observable variables; for example, Attanasio and Davis [12] find that the average consumption of high school graduates moved together with their average real wage in the 1980s. In principle, such a fluctuation specific to a group of consumers can be reduced by having a contract contingent on publicly observable variables, and it is hard to imagine that the problem of moral hazard keeps such a contract from being
made in markets.

Following the theoretical framework of Thomas and Worrall [111], Hayashi [57] presents a simple model of consumption allocation with self-enforcing constraints. In this model, the self-insurance opportunity is ruled out by assumption, and, consequently, consumers can rely on only the self-enforcing insurance contract in order to insure idiosyncratic shocks. This model generates the following empirical implication.

In complete markets, the Lagrange multiplier $\lambda_i$ in equation (1) is constant over time; the constancy of $\lambda_i$ suggests that agent $i$ commits to the contracts made at time 0. When a self-enforcing constraint is binding, on the other hand, the Lagrange multiplier is adjusted\textsuperscript{17}. Once the Lagrange multiplier is adjusted for a fraction of consumers, tight implications of complete markets such as equations (2) and (3) break down immediately for the sample including constrained and unconstrained consumers. In simulation, Hayashi [57] finds a weak, but statistically significant correlation between individual income and consumption. Again, the empirical relevance of the above implication may depend on how much the above self-enforcing insurance improves on self-insurance in terms of efficiency gains. A reasonable efficiency gain justifies a serious look at the above self-enforcing insurance contract.

3.3. Disproportionate allocation of aggregate shocks

Even in the case of complete markets, aggregate shocks are allocated disproportionately among consumers. Less-risk-averse consumers take more aggregate risks than more-risk-averse consumers, while attitudes toward risks may differ between the young and the old in the overlapping generations setup. Dumas [37], for example, analyzes the financial trading between risk-takers (less-risk-averse) and hedgers (more-risk-averse) in dynamically complete markets.

\textsuperscript{17}In complete markets, a planner assigns an inverse of the Lagrange multiplier for individual budgets as a weight for individual utility (see Huang and Litzenberger [67]). Precisely stating, the inverse of the Lagrange multiplier is adjusted in the binding case, since the model of Hayashi [57] is constructed as the problem of a planner.
Heterogenous attitudes toward risks, however, do not necessarily have a significant impact on the construction of representative agents. Some aggregation properties of complete markets still survive despite the heterogeneity in risk aversion (see Huang and Litzenberger [67]), and heterogeneous preferences may not be responsible for the empirical failure of representative agent models.

In this subsection, we discuss the case where market incompleteness makes aggregate shocks distributed disproportionately among investors, and see how such a disproportionate allocation has an impact on the representative agent representation. In particular, to match the observation that only a fraction of the entire population hold stocks in portfolios (see Section 2.3), we are interested in the limited participation in stock markets. While limited participation is an exogenous constraint in the models discussed below, the exclusion from stock markets may arise due to high transaction costs, expensive information costs, small sizes of wealth, low educational backgrounds, ignorance of stock markets practice, and so on. Saito [102] and Basak and Cuoco [16] present analytical models with limited participation as exogenous constraints.

In Saito [102], the stockholders satisfy the Euler equation with respect to both stock returns ($R_s(t, t+1)$) and safe bond returns ($R_f(t, t+1)$), while the non-stockholders satisfy with respect to only safe bond returns. In this case, aggregate consumption only partially reflects the consumption of investors who are facing complete markets (the consumption of the stockholders), and the Euler equation defined in terms of aggregate consumption is misleading. Saito [102] shows that the Euler equation holds for neither stock returns nor safe bond returns at the aggregate level. More precisely, the Euler inequalities hold at the aggregate level as follows:

$$\frac{1}{1 + \rho} E_t \left( R_f(t, t+1) \frac{u'(C(t + 1))}{u'(C(t))} \right) < 0,$$
and

\[
\frac{1}{1 + \rho} E_t \left( (R_s(t, t + 1) - R_f(t, t + 1)) \frac{u'(C(t + 1))}{u'(C(t))} \right) > 0.
\]

A brief explanation for the above inequalities is offered as follows. First, the non-stockholders invest the entire wealth in safe bonds due to participation constraints, and they create more demand for safe assets in this case than in the case without any participation constraints (the case of complete markets). Consequently, risk-free returns go down, and the former inequality obtains. Second, the non-stockholders cannot have access to stock markets, and their Euler equation with respect to excess returns on equities does not hold. The latter inequality at the aggregate level reflects the failure of the Euler equation of the non-stockholders.

These Euler inequalities imply that a risk-free rate is lower and an equity premium is larger in the case of limited participation than in complete markets. The more wealth (the wealth held by the non-stockholders) is excluded from stock markets, the more prominent is the degree of the above two inequalities. Restating this, the asset pricing of this model depends on the wealth distribution between the stockholders and the non-stockholders. In addition, under the above setup, the forecast error of the Euler equation defined in terms of aggregate consumption has predictable components which are pro-cyclical in the equation for risk-free rates, and counter-cyclical in the equation for equity premia. These predictions in regard to the Euler equations are broadly consistent with the empirical finding discussed in Section 2.

Devereux and Saito [35] apply a similar idea to a two-country model. In their model, the wealth of one country is excluded from the stock market of the other country, and

---

18 Notice that the Euler equation with respect to excess returns holds for the consumers in the case of Grossman and Shiller [45]. The reason for this is that in their case, insurance for idiosyncratic shocks is not available, but consumers are still able to have access to any financial assets including securities.

19 Baxter and Crucini [17] construct a similar two-country model in discrete-time, and solve it numerically.
only non-contingent claims (risk-free assets) are traded between the two countries. As in Saito [102], the return in the international bond market changes over time depending on the cross-country wealth distribution.

Among the above models of Saito [102], Basak and Cuoco [16], and Devereux and Saito [35], the wealth distribution between heterogenous agents changes in response to aggregate shocks (or country-specific shocks), and asset pricing depends on the evolving wealth distribution; the dependence of asset pricing on wealth distribution never appears in representative agent models. Because of this feature, tracing the changing wealth distribution is fairly important for characterizing the dynamics of asset pricing. Saito [103] presents a convenient analytical tool to track the evolution of wealth distribution in a continuous-time setup.

4. Welfare implications The theoretical research into incomplete markets discussed in the previous section triggers a reassessment of welfare implications for dynamic resource allocation. In particular, business cycles, capital accumulation, economic growth, and international risk-sharing have been evaluated in terms of efficiency in the context of incomplete markets. This section briefly surveys the theoretical research in regard to the welfare evaluation of dynamic allocation.

**costs of business cycles** Lucas [82] presents a provocative argument that the cost of business cycles is extremely small, and that welfare gains from eliminating the fluctuation of aggregate consumption is almost negligible. He evaluates the welfare cost of business cycles using representative agent models. The literature on incomplete markets reconsiders this striking argument from the perspective of asset pricing.

The evaluation of costs of business cycles is indeed intimately connected with the pricing of aggregate risks; business cycles are costly for consumers who demand a large risk
premium in exchange for taking aggregate risks. Accordingly, the plausibility of the representative agent framework for evaluating the cost of business cycles has close relations to that for pricing aggregate risks. A deep doubt on representative agent models as asset pricing models immediately raises a serious question as to how effective they are for evaluating costs caused by aggregate fluctuations.

Campbell and Cochrane [20] build a representative agent model with habit formation such that large equity premia may be yielded. They show that the implied cost of business cycles in their model is much higher than that calculated by Lucas [82]. Their finding indicates that models able to account for the equity premium puzzle may bring about the prediction of costly business cycles.

As discussed in Section 3, models of incomplete insurance with borrowing constraints may yield a reasonable magnitude of equity premia. Given a close connection between asset pricing and welfare evaluation, the welfare assessment of business cycles may be sensitive to whether markets are complete or incomplete. Imrohoroglu [69] demonstrates that the more severely borrowing constraints are binding, the larger is the cost of business cycles. Adopting the framework of Constantinides and Duffie [29], Heaton [61] finds both large equity premia and costly business cycles in the case where the variance of permanent idiosyncratic shocks is counter-cyclical. These models suggest that those who fail to diversify idiosyncratic shocks tend to evaluate the cost of business cycles highly. In other words, the cost of business cycles may be larger in incomplete markets than in complete markets.

Atkeson and Phelan [9] challenge the above view. Using a simple model of employment risk, they present the case where, in reaction to the reduction in aggregate fluctuations, economic agents choose to face more idiosyncratic employment risk. Restating this case, in spite of reducing aggregate risk by economic policy, the total amount of risk faced by individual agents is almost the same as before. Given this endogenous determination of
idiosyncratic risk, the cost of business cycles is smaller in complete markets where idiosyn-
cratic shocks are insured, than in incomplete markets where idiosyncratic shocks remain
uninsured. If it is the case, the welfare evaluation based on representative agent models
puts the least upper-bound on the cost, and basically supports the methodology of Lu-
cas [82]. The plausibility of the above Atkeson and Phelan model must be examined by
empirical research on how idiosyncratic risk responds to aggregate risk.

**welfare assessment of capital accumulation and economic growth**  As men-
tioned before, incomplete insurance usually induces precautionary savings\(^{20}\). Such precau-
tionary savings have effects on capital accumulation in neo-classical growth models, and on
economic growth in endogenous growth models. Given the welfare impact of market incom-
pleteness, some studies explore a theoretical relationship among incompleteness, welfare,
and economic growth (or capital accumulation). A welfare assessment of capital accumu-
lation and economic growth often differs substantially between complete and incomplete
markets.

*(neo-classical growth)* In the context of neo-classical growth models without any
aggregate shocks, following Bewley [19], Aiyagari [4] demonstrates that net interest rates
(net-of-depreciation interest rates) are lower than the rate of time preference when income
insurance is incomplete. Aggregate physical capital is, therefore, over-accumulated in the
sense that the level of capital is above the modified golden rule. In this case, the taxation
of capital income prevents the over-accumulation of capital, thereby raising economic wel-
fare. Restating this, the optimal tax rate on capital income is positive in the Bewley type
stationary economy. The above result is strikingly in contrast with Chamley [23] where the
long-run optimal tax rate on capital income is zero in full insurance economies.

\(^{20}\)A theoretical characterization of precautionary savings is given by Kimball [73] and others.
(endogenous growth) The most important virtue of endogenous growth models is that, unlike in neo-classical growth, economic growth reflects not only exogenous technological factors, but also other factors including preference, innovation processes, government policy, and market structure. In the literature, market incompleteness has been regarded as one of the most responsible factors. In particular, they are interested in how financial development affects economic growth. Here, financial development is defined as a process in which financial markets are being completed.

For example, Devereux and Smith [34], Obstfeld [93], Jitsuchon and Saito [71] present endogenous growth models with incompleteness as exogenous constraints. Bencivenga and Smith [18] introduce financial intermediation into an endogenous growth model, while Greenwood and Jovanovic [44] and Saint-Paul [106] construct models where economic growth and financial development are jointly determined.

As summarized in Pagano [96], there are two opposite effects of financial development on economic growth among the above models. As a consequence of risk diversification, on the one hand, physical capital may be invested on riskier, but more productive opportunities. Savings may, on the other hand, decline due to weaker precautionary motives. The former effect enhances economic growth, while the latter lowers it. Which effect dominates depends on underlying structural parameters.

The following example illuminates how the above opposite effects work. In Jitsuchon and Saito [71], the average aggregate consumption growth \( E \frac{\Delta C}{C} \) is characterized by

\[
E \frac{\Delta C}{C} = \epsilon (E R_f(t) - \rho) + \frac{(\gamma + \gamma \epsilon) \sigma^2_h}{2},
\]

where \( \epsilon \) implies the elasticity of intertemporal substitution, \( \gamma \) is the degree of relative risk.

\[^{21}\text{Endogenizing incompleteness from the problem of moral hazard, Tsidon [117] and Aghion and Bolton [2] present models where capital accumulation and financial development are jointly determined.}\]
aversion, $\sigma_h^2$ is the variance of uninsured permanent shocks, $\rho$ is the rate of time preference, and $E R_f(t)$ is the average risk-free rate. Notice that consumption growth is equal to economic growth in endogenous growth models.

This equation implies that given asset returns ($E R_f(t)$), the economy grows faster in the presence of uninsured idiosyncratic shocks than in the absence of such shocks. The main reason for this is that uninsured idiosyncratic shocks motivate precautionary savings, and promotes capital accumulation. Indeed, the coefficient on the uninsured component in the second term ($\gamma + \gamma e$) implies the degree of relative prudence defined in the non-expected utility (see Kimball and Weil [74]).

In the above equation, the first term of the right hand side corresponds to the standard $Ak$ type endogenous growth model (see Rebelo [98]), while the second term represents the precautionary saving motive. When financial development provides risk-diversification opportunities thereby lowering $\sigma_h^2$, $E R_f(t)$ in the first term increases, while the second term decreases due to less precautionary savings. Jitsuchon and Saito [71] show that the latter effect dominates when $\gamma$ is greater than one. In a related context, Jappelli and Pagano [70] introduce liquidity constraints into an endogenous growth model. When credit markets are developed in their model, less binding constraints may lead to lower saving rates, thereby reducing economic growth.

The above predictions all suggest that financial development may have a negative effect on growth. In most cases, on the other hand, completing financial markets improves economic welfare. While financial markets are being developed, therefore, welfare and growth may move in an opposite direction. In other words, faster growth may not bring about higher welfare. Interestingly enough, Jappelli and Pagano [70] point out that a deep interest of the public in high economic growth sometimes work in order to repress insurance and consumer credit markets in several ways; such repression of financial markets maintains
economic growth at the expense of social welfare.

**welfare gains from international risk-sharing** As discussed in Section 2.4, the consumption correlation puzzle indicates poor risk-sharing among countries, while the home bias puzzle suggests that there is a distinct lack of international diversification in financial markets.

Obstfeld [93] and van-Wincoop [118] indeed demonstrate that each country could receive substantial gains from international risk-sharing. In their welfare calculation, they compare the outcome of perfectly completing international markets with that of assuming each country as an autarkic economy. Since there is an obvious welfare gain from a shift from autarky to full insurance, the only issue is how much is the gain from full insurance.

Such a simple welfare comparison, however, may involve some problems. First of all, as repeatedly discussed so far, there is market incompleteness in a domestic country itself. Therefore, potential opportunities for risk-sharing may exist not only in international markets, but also in domestic markets. Tesar [109] finds that, allowing for self-insurance activities in each country, there are negligible gains from international risk-sharing. She suggests that such small gains are not sufficient to offset transaction costs incurred in international markets, thereby justifying the lack of global risk-sharing.

What is more fundamental, regarding complete isolation from international markets as status quo seems too simple an assumption. Developed countries usually have at least partial access to international financial markets. In particular, it is well-known that money markets (short-term bond markets) are integrated globally. A more accurate description should be, at status quo, countries are exposed to some extent to international intertemporal prices.

In the above case, opening another market generally has an impact on existing prices,

---

22 Cole and Obstfeld [26] present a more conservative estimate of welfare gains.
and generates complicated effects on welfare distribution among agents. Opening markets, therefore, may not result in Pareto-improving; someone may be worse-off due to market opening. Such a general equilibrium impact on welfare distribution has been noted in the general equilibrium literature. Furthermore, Hart [53] and Newbery and Stiglitz [92] present even stronger results that opening another market makes everyone worse off.

Using a two-country setup, Devereux and Saito [35] compare the case where only non-contingent claims are traded internationally (the bond regime) with the case where markets are globally complete (full risk-sharing). There are a debtor country and a creditor country in the bond regime. When the international economy shifts from the bond regime to full risk-sharing, precautionary savings in each country decrease as a result of diversifying country-specific shocks, thereby raising bond returns in international markets. Then, a country which used to be a debtor loses the advantage of low interest rates. For this debtor country, the regime shift sometimes makes the benefit of risk-sharing dominated by the cost of losing the opportunity of low interest rates. Devereux and Saito [35] argue that such potential effects of financial integration on the cross-country welfare distribution may make some countries stick to status quo where risk sharing is only partial.

5. Conclusion The research into incomplete markets has been developed in order to account for asset pricing anomalies and heterogeneous consumption behavior, both of which are contradictory to the predictions of the complete markets framework. This survey has reviewed such development in empirical and theoretical respects. To conclude this survey, the author offers two remarks on future research.

First, the literature has so far put more emphasis on a positive side of incomplete markets models. The main reason for this is that the observed deviation from the complete markets paradigm is fairly striking; explaining the dramatic deviation itself has been intellectually exciting. In future, a research emphasis may be shifted from a positive side
to a normative side. A complete markets economy has well-known desirable welfare features. The substantial deviation from such an economy tells us that economic agents suffer significant welfare losses. How much are their welfare losses? How are such welfare losses remedied? Which policy results in welfare improvement? How much does a financial innovation raise welfare? How does a financial integration have an impact on welfare distribution among concerned countries? Theoretical models will be constructed in order not only to account for empirical findings, but also to answer these questions.

Second, to answer the above questions more fundamentally requires theoretical models which are able to endogenize market incompleteness from first principles. As discussed in Section 3, two different endogenous constraints may yield seemingly similar positive (empirical) predictions for asset pricing, but the choice of policy depends on which fundamental factor underlies these constraints. Without knowing true economic reasons, policy makers cannot write policy prescriptions properly.
REFERENCES


Dynamic Allocation and Pricing in Incomplete Markets: A Survey


