Computer Assisted Portfolio Selection
Theory and Practice

Harry M. Markowitz*

Background

In 1999 Sherrie Grabot, CEO of GuidedChoice (GC), flew down from San Jose to meet with me in the Harry Markowitz Company conference room, to discuss her strategy for entering the 401(k) advisory industry using a different business model than anyone else used then. I expressed great interest in her project, gave her a copy of *Individual Versus Institutional Investing* (Markowitz 1991), and explained that it outlined a "Game of Life" computer simulation that included a family's health, housing, educational achievements and plans, skill sets, social security, insurance, etc., in addition to investment decisions. I did not propose that GuidedChoice, or anyone else for that matter, create a complete Game-of-Life model. Rather I proposed it as an ideal, a North Star, towards which we would direct our model-building, starting with our immediate objective of helping investors save for their retirement.

On the product development side, Sherrie said that she had a team in India developing a www-based interactive front-end. She asked if I could form a team to develop the technical back-end. I told her that I would try to hire two former colleagues, Ming Yee Wang (who has a Master's degree in math from the Courant Institute) and Gan Lin Xu (who has a Ph.D. in math from Carnegie Mellon). If I could succeed in that, and if she were willing to locate the technical office in San Diego, then she had her technical team. Ming and Gan Lin did in fact join, and the GC San Diego office began in January of 2000.

Markowitz (1991) was not the first time I expressed ideas pointing in the direction sketched above. Sections of Markowitz (1959) Chapter 13 that discuss non-portfolio decisions that impinge on, and perhaps must be made jointly with, portfolio selection had similar motivation. In effect, Markowitz (1991) spells out ideas in Chapter 13 of Markowitz (1959).

The responsibilities of the GC San Diego technical team included:

> (1) The generation of a mean-variance frontier at an asset class level, based on forward looking estimates of means, variances and covariances, and constraints on the choice of asset class portfolios.
>
> (2) The development of software to choose, for any given 401(k) plan and for each asset class portfolio, a portfolio using only those investment companies permitted in the particular client's plan.

(3) The development of a Monte Carlo model to generate a probability distribution of how much the participant will be able to spend per month after retirement, taking into account the participant's savings rate, the company's matching policy, the asset class mix chosen from the efficient frontier, and hundreds of randomly-drawn asset-class return scenarios.

The San Diego team delivered almost on time. The team developing the web-based front-end did not. We waited, and waited, and waited. The developers of the front-end finally delivered, but their software bombed-out when Sherrie demonstrated the product to her investors. On further investigation it became clear that the front-end was a bug-filled hopeless mess.

Sherrie asked the San Diego team to try to supervise the debugging or rewriting of the front-end. But none of the three of us had internet programming experience. I, for one, had learned many programming languages in my time—including FORTRANs 1 and 2, PL/I, C, C++, Java, and had designed and supervised the programming of the SIMSCRIPT programming language. But I could not make heads or tails out of database-supported, interactive internet programming with its arcane multi-vendor requirements.

We tried using consultants and an internet-knowledgeable programmer hired for us by a manager in Sherrie's San Jose office. But the consultants were ineffective in the given situation, and the programmer was incapable of making progress single-handedly, without the support of a large team and strong guidance from the top. It was painfully obvious that debugging the front-end, or reprogramming that entire part of the project in the time-frame necessary with the resources Sherrie could then muster, was beyond the capacity of any normal human team.

Meanwhile Gan Lin quietly taught himself database supported internet programming, and pulled the bugs out of the front-end. Sherrie now was within closing-distance of her first product, GuidedSavings. There was still much to be done by both the San Jose and San Diego teams to make the product work as Sherrie wanted it to, rather than as it had been built orginally, and to create a user-friendly front-end.

It is now over thirteen years since the GC San Diego technical team started. Since then Sherrie's team has successfully marketed, and the San Diego team supported, the GuidedSavings product that guides participants' savings until they retire. We have also rolled out a second major product, GuidedSpending. This can be used before retirement, but continues into retirement until the participant's inevitable end. Major product design decisions for these products have always been the result of discussions until a consensus was reached by the original technical team. I will now refer to the three of us as the San Diego R&D committee as distinguished from our GC technical support staff. Ming is the R&D committee's leader. He is in regular touch with Sherrie, and sets the R&D committee's agenda as to what we should work on first and what later. He also maintains all models and software not related to database and internet. The latter is the responsibility of a group under Gan Lin. In addition, for over a year now Tom Anachini has been a member of the R&D committee. He has taken over from Ming most research analysis, such as the Principal Components Analysis and investigations of fixed income models described below.

Over the years the R&D committee has met at least once almost every week and has had no end of problems to solve. These included the initial GuidedSavings design according to Sherrie's specifications; unanticipated complexities when the product was used for real; hot new product opportunities brought to us by marketing, almost none of which panned out (marketing, now under David Bernard, is much more disciplined); the need to re-estimate means, variances and covariances; and the design of our latest product, GuidedSpending. Below I describe three of the R&D committee's problem areas to illustrate our on-going struggle with the theory of practice and the practice of theory.

Database Design

As recounted in the previous section, when the team originally assigned to build the front-end failed, Gan Lin learned internet programming and debugged the code. My job, naturally, was to understand what the programmers had done, make sure that that was what Sherrie Grabot wanted, and help steer the team toward Sherrie's desired goal. The original developers had essentially used the ER (Entity Relationship) view of system description. As shown in Markowitz (1983) this is logically equivalent to my own favorite view, the EAS (Entity, Attribute and Set) view. In practice, a major difference between the two is the method of documentation. The original developers had documented the structure of their database using the "Rational Rose" software. Entity-types (a.k.a. "object classes") were represented by circles; relationships were represented by lines connecting these circles. The overall picture may best be visualized as a large platter of meatballs and spaghetti. In particular, it was often difficult to trace connections between meatballs on one side of the platter and those on the other side. Compared with the EAS table described below, the Rational Rose diagram was incomplete and a thorough mess.

The EAS view of status description is part of the EAS-E (Entities, Attributes, Sets and Events) view of dynamic systems modeling. This was first used in the original SIMSCRIPT programming language (Markowitz, Hausner and

Karr 1963), now called SIMSCRIPT I to distinguish it from SIMSCRIPTs II and III that followed. SIMSCRIPT I was conceived as a simulation language, hence its name. In contrast, SIMSCRIPT II was conceived as a general purpose programming language (See Markowitz 1979). It was documented and, to a certain extent implemented, as a series of levels. Level 1 was a simple expository language; Level 2 brought the language up to approximately a FORTRAN 2 level; and Level 3 brought it up to roughly an ALGOL level. Level 4 introduced Entities, Attributes and Sets. Level 5 introduced Events and various simulation-oriented matters such as facilities for accumulating performance statistics, and random number generators with commonly used probability distributions.

Level 6 introduced database entities into the language. The idea here was that the basic SIMSCRIPT premise—that the world consists of Entities, Attributes and Sets—applies to “the world” as represented by a database as well as that in a simulation model.

Level 7 was planned to make available to the systems programmer the Language Writing Language (LWL) in which SIMSCRIPT II was itself programmed. In the first instance SIMSCRIPT II was bootstrapped from SIMSCRIPT I. But thereafter versions of SIMSCRIPT II for new computers or operating systems were programmed using SIMSCRIPT II’s EAS description of itself.

It is not necessary to program in SIMSCRIPT to make use of the EAS-E worldview. For example, Jacobs, Levy and Markowitz (2004) presents the JLMSim stock-market simulator. This includes entity-types such as Investor, Portfolio Analyst, Statistician, Security, Trader, and Order. A Security has attributes such as Last_trade_price and Volume_so_far_today. It "owns" a set of Buy_orders and a set of Sell_orders, each of which is sorted by bid or asked price. JLMSim Events include time-to-reoptimize, time-to-review-order-status, etc. JLMSim was conceived in EAS-E terms, and then implemented in C++.

Exhibit 1 here presents part of the EAS table used to document and implement the GuidedChoice database structure. I prepared a first draft based on the Indian team's Rational Rose diagram. The scope of the resulting EAS table was extended considerably, based on the R&D committee's many discussions with Sherrie as to how the product would work, and also based on Gan Lin's detailed questions as he began to implement the database facility.

The first column of Exhibit 1 indicates "Entity-type," such as Person, Dependent and Planned_disbursement. Individuals of a given entity-type are referred to as "entities" or "individuals." The second column contains the names of attributes of entities of the just listed entity-type. Examples of attributes include Birth_date, Soc_sec_num and Marital_status of Person. The third column of the exhibit lists sets "Owned By"—that is, sets *associated with*—individuals of the particular entity-type. For example, each Person owns sets called Dependents, Portfolios and Planned_disbursements (to the person). The fourth column of the exhibit indicates the data-type of an attribute or the entity-type of the members of a set. For example, the Birth_date of Person is in the Date format, and the members of the set called Dependents are of entity-type Dependent.

The fifth column is labeled precision. It indicates, for example, that 32 characters are allotted for a State's name, which is also assigned a two character State code. The final field of the exhibit includes comments and cross-references. The list of entity-types under the comments/cross references heading across from entity-type Authorization_memo, for example, indicates that individuals of this entity-type are referenced by The_SYSTEM and entities of entity-type Person. Such cross-references are not a standard part of an EAS table, but Gan Lin requested them to facilitate implementation.

An EAS description invariably has an entity-type called The System. This represents the "system as a whole." Exhibit 1 shows that the GC System owns various sets of "top level" entity-types. Entities of other entity-types are accessed through these top level entities. Top level entity-types include everything from Sponsors (of 401k plans), Record keepers, Securities, and so on down to entries in the Transaction_log_book, Event_log_book and Error_log_book. Attributes of The_System include Federal limits on the dollars or percent that may be contributed to a 401(k) plan. No attempt has been made to update comments such as the then-current limits on permitted contribution size.

As the exhibit illustrates, it is often useful to include what SIMSCRIPT refers to as "compound entities," and what mathematicians call Cartesian products. This extract from the full GuidedChoice EAS table includes, as its only compound entity, a Person_X_Comp_type combination (such as the overtime pay of a participant as opposed to his or her base salary). The full GuidedChoice EAS documentation has many more examples of compound entities, as does the extract of the JLMSim EAS table in Jacobs, Levy and Markowitz.

The leader of the original front-end programming team asserted that the Rational Rose diagram was the only documentation of their system other than comments in the program itself. I do not know to what extent their lack of a well-documented plan contributed to their failure. But it certainly didn't help.

Utility Functions

Both GuidedSavings and GuidedSpending offer initial advice based on the Monte Carlo evaluation of various proposals. (Participants are encouraged to vary the parameters of this initial advice and review the resulting Monte Carlo analysis as often as they wish.) To form our initial advice, a utility function was needed for each of our two products. These are used to assign a utility number to each run of the Monte Carlo analysis and, from these, an estimate of the expected utility of a proposal. In the case of GuidedSavings, utility had to be a function of wealth at retirement time:

$$U = U(W_T) \tag{1}$$

For GuidedSpending utility had to be a function of the participant's simulated consumption stream plus bequest amount

$$U = U(C_1, C_2, ..., C_T, W_{T+1}) \tag{2}$$

We will discuss each of these in turn.

(1) *GuidedSavings.* As Mossin (1968) and Samuelson (1969) have shown, if $U(W_T)$ in Equation (1) were logarithmic or a power function, and certain other conditions were true, then the utility function would be myopic. In other words, if investment opportunities did not change, the optimal asset allocation would not change as the participant approaches retirement. Since it seems plausible that the participant should usually shift towards bonds as retirement approaches, alternatives are frequently sought.

Certain life-cycle models assume that the present value of one's future labor is like a bond. As this implicit bond shrinks in value the participant is advised to shift from stocks to bonds, thus effectively maintaining a fixed stock-bond ratio. The problem with this argument is that many investors probably consider their employment incomes more like equity returns than fixed income. Examples of such include everyone who was concerned about their job, or at least their take-home pay, during the great recession. About the only individuals who might consider their regular salary to be "fixed" are tenured (associate or full) professors. Thus the life cycle plans in question may be thought of as "of tenured professors, by tenured professors, for tenured professors."

A more plausible explanation for our intuition that the investor should become more cautious as retirement approaches has to do with the opportunity for growth in the long run when much time remains versus the seriousness of substantial losses if little time remains. We found that, with suitable choice of parameters, this motivation and its effect can be adequately represented by the simple device of having some aspiration level W_0 for retirement real wealth, and letting

$$U = \begin{cases} \beta_1 W_T^{\alpha_1} & \text{for } W_T \geq W_0 \\ \beta_2 W_T^{\alpha_2} & \text{for } W_T \leq W_0 \end{cases} \qquad (3)$$

where $\alpha_2 < \alpha_1 < 0$, and β_1 and β_2 are chosen so that

$$\beta_1 W_0^{\alpha_1} = \beta_2 W_0^{\alpha_2} \qquad (4)$$

(2) *GuidedSpending*. Our utility function of form (2) uses two numbers available in GuidedSpending but not in GuidedSavings. We ask the GuidedSpending user for two levels of consumption, C_U and C_L. At any point in simulated time in a Monte Carlo run, a proposed current consumption C is determined by an actuarial calculation that allows for the participant living somewhat longer than expected. If C exceeds C_U the difference, $C - C_U$, is saved. If C is less than C_L then $C_L - C$ is dissaved—if available. As described below, C_U and C_L are used in the calculation of U, as is an aspiration level of bequest B. If the participant declines to supply either B or C_L and C_U, then default values are supplied at levels dependent on the user's likely retirement wealth.

To compute the U attached to a particular consumption history and bequest, we form a score S by combining the history's average consumption level A and its maximum year-to-year decline in consumption D:

$$S = A - \alpha D$$

The idea here is that it is better to start poor and end rich than vice versa. We then normalize S forming a "normalized score," NS, such that if $A = C_U$ and D = 0 then

$U = 1$ whereas if A = C_L and D = 0 then $U = 0$. A term reflecting W_{T+1} versus B is added for a final score FS. Finally, utility is computed as a function of FS

$$U = f(FS)$$

where f is a smooth curve with $U = 1$ as an asymptotic upper bound and with U dropping off steeply as FS drops below zero.

Principal Components Analysis (PCA)

We performed a PCA to model the covariance structure of asset class returns. This is plausible for its own sake and is especially important because our asset-class data series are of differing length. Since PCA is not uncommon, I will only report here ways in which our use of PCA may be unusual if not unique.

(1) From the loadings of different asset classes on the PCs, it was clear that PCs 4 and 5 had to do with fixed income. Specifically, PC 4 had to do with changes in the level of the yield curve whereas PC 5 had to do with its slope. Since we have a separate model for generating short term versus medium term versus long term rates, we combined PCs 4 and 5 and refer to it as PC 4.5.

Because of the current historically low interest rates, and the probability of these rising over time, we distinguished between near-term expected total returns to fixed income securities as compared with long-term estimates, and considered the probable paths from the one to the other. Over many weeks Tom Anachini explored, and the R&D committee discussed, alternative interest rate models. We converged to certain parameter settings for the CIR (Cox, Ingersoll and Ross 1985) model. We also found that i.i.d. draws from a normal or lognormal distribution for successive values of PCs would be quite unrealistic. Rather, PCs have had (and we assume they will continue to have, but we will monitor this) an autoregressive structure with Pearson Type IV return distributions.

Now

Except for Sherrie and a helper, who will remain in the Palo Alto area for the time being, GuidedChoice now resides in its new San Diego offices, with the technical folks occupying one wing, and the marketing and customer support folks occupying another. Ming, Gan Lin, and Tom have their above-noted responsibilities, as well as participating in the weekly R&D committee meeting. It seems my role is "Senior Theoretician." There is one more responsibility I enjoy taking on from time-to-time. Whenever someone who does not have a primarily clerical job joins our staff, I schedule a series of weekly meetings with them to explain the EAS structure of GuidedChoice, part of which is in Exhibit 1 here.

Note that I speak of this as GuidedChoice's EAS structure, not just its database's structure. After all, its database is what GuidedChoice officially remembers of itself and the world. As to other, unofficial, memories, no one seems to mind if we break from examining EAS structure for a few minutes for me to tell tales of the "old days," such as, for example, how the San Diego part of the GC story started in the Harry Markowitz Company conference room, and how—when GuidedChoice seemed doomed—Gan Lin saved the day.

Other aspects of GuidedChoice

My discussion above of the database design, utility functions and principal components analysis used by GuidedChoice reflect my own interest in new or ongoing challenges to financial theory in support of financial practice. GuidedChoice also makes use of "old stuff" that has long been standard in the broad "MPT Industry" of which the 401(k) advisory services sector is a relatively small part. The remainder of this paper is devoted to a brief top-down view of GuidedChoice as a whole, and a briefer look at the MPT industry and the 401(k) advisory services' place in it.

GuidedChoice is divided into three parts: (a) product development and maintenance under Ming Wang, (B) marketing and client support under Dave Bernard, and (C) administration. The heads of (A) and (B) report to Sherrie Grabot, who also supervises (C). As I noted earlier, the group under Ming Wang had as its main responsibilities the development and use of software to:

> (1) Generate mean-variance frontiers at an asset class level;
>
> (2) Select, for any given 401(k) plan and each asset-class portfolio, a portfolio using only those investment companies permitted by a particular plan; and
>
> (3) Use Monte Carlo simulation to generate a probability distribution of how much a participant will be able to spend per month after retirement.

To this was added the responsibility to:

> (4) Develop (as well as maintain and run) the interactive front end, the GuidedChoice database, and related activities such as periodic portfolio rebalancing and performance reporting.

(1) The first of these is definitely “old stuff” for the MPT industry. The idea of an efficient frontier was proposed in Markowitz (1952), while the idea of applying it in a top down manner is, I assume, a by-product of discussions that followed the Brinson, Hood and Beebower (1986) assertion that most of professional portfolio managers’ performance was due to asset class selection.

(2) Sherrie Grabot’s original idea was to offer a client *managed accounts* rather than *advice only* in the manner of *Financial Engines*, the firm that Bill Sharpe created (thereby creating the 401k advisory service industry). Offering of managed accounts implied a need for a method of assigning—to each asset-class portfolio—a portfolio comprised of those investment companies offered by a given plan. This involves a tradeoff between an estimate of performance versus tracking error. Since the former is linear and the latter is quadratic, we have here essentially a mean-variance side-calculation for each plan and asset-class portfolio. Ming Wang led the research in this area, as well as programming the algorithm. A principal result of Ming’s research is that, as others have found, the largest determinant of net performance is management fees and portfolio turnover.

(3) Following the precedent Bill Sharpe set in Financial Engines, we built a Monte Carlo simulator for GuidedSavings that generates probability distributions of wealth at retirement time. We convert these into a probability distribution of reasonable post-retirement real consumption expenditures. For GuidedSpending

we generate consumption and bequest patterns as discussed in a prior section. These calculations take into account a participant's savings rate, their company's matching policy, the asset-class mix chosen from the efficient frontier, and hundreds of randomly-drawn asset-class return scenarios. A frequent consequence of a participant's use of GuidedChoice's interactive Monte Carlo-based system is for the participant to increase his or her savings rate. Remarkably, it was found that participants who used GuidedSavings increased their savings rates by an average of 110%.

(4) We distinguish "offline" and "online" actions by GuidedChoice software. Computing an asset-class efficient frontier is done periodically, offline. Computing an investable portfolio corresponding to each of seven asset-class portfolios using permitted investment companies is done offline, when a new client (i.e., plan sponsor company) joins us, or when the client changes its investable-security list.

The regular "online" cycle, including day-to-day and periodic actions, includes new participants enrolling using our interactive software; GuidedChoice placing orders with clients' record keepers to execute the participant's savings and investment decisions; periodically considering whether rebalancing trigger-levels have been reached and, if so, placing rebalancing orders with record keepers; and periodically producing status reports for participants and summary reports for the client and our own monitoring.

The technical team also works closely with the marketing and client support team as needed. In the development of a new product, or addition of a new feature to an old product, it is up to the technical team to decide what should be the inputs, outputs, and calculations required to properly serve the participant; the client support team designs the interactive interface to be used by the participant; and the technical team implements this design. But the process is much less linear than just outlined, with much energetic interchange between the technical team, the marketing/client support team, and Sherrie Grabot before a consensus emerges.

From time to time the technical team also does special analyses for the marketing support folk. For example, Figure 1 is based on an analysis by Tom Anachini of the technical team, at the request of one of GuidedChoice's larger clients as conveyed by the client support team. The figure shows the annualized return (since inception by the client) to: (a) all 401(k) participants, (b) those who did not use GuidedChoices service and (c) those who did. The difference in results is spectacular. The median level of annualized return to accounts managed by GuidedChoice was roughly two percentage points (200bps) greater than those we did not manage. (Variations within each category are due in part to differences in start and end times for participant's participation). More generally, by collecting participant individual rate of return data from several record keepers and plan clients, GuidedChoice found that their managed accounts outperformed non-managed accounts by 1% to 3% annually. There was also a sizable reduction in performance variability.

The entire sector that provides managed accounts for 401(k) and other defined contribution (DC) plans has grown fast and at an accelerating rate in its brief history, but there is still much room for further growth. “By the end of 2012, the top eight managed account firms had $107.9 billion in assets under management, less than 2% of the total $6.92 trillion in U.S. DC assets, according to Cerulli Associates, Boston.”[1] With outperformances such as illustrated in Figure 1, it seems plausible to expect that the 401(k) advisory service sector, in particular, and the DC managed account sector in general will rapidly increase its fraction of the entire DC sector.

Endnotes

*The recollections recounted, and the views expressed, here are those of the author and do not necessarily reflect those of GuidedChoice, Inc. or any individual therein besides myself.

[1] "Interest in managed accounts is low. Consultants, plan execs, participants appear hesitant to implement as QDIA". Robert Steyer, July 2013. http://www.pionline .com/article/20130722/PRINTSUB/3079980.

References

Brinson, Gary P., L. Randolph Hood and Gilbert L. Beebower (1986), *Determinants of Portfolio Performance*, Financial Analysts Journal 42 (4), pp. 39-44.

Cox, J. C., J. E. Ingersoll Jr., and S. A. Ross, 1985, “A Theory of the Term Structure of Interest Rates,” Econometrica, 53, 385-407.

Jacobs, B. I., K. N. Levy, and H. M. Markowitz (2004) *Financial Market Simulation*, Journal of Portfolio Management (30th Anniversary) pp. 142-152.

Markowitz, H.M. (1959), “Portfolio Selection: Efficient Diversification of Investments,” Wiley, Yale University Press, 1970, 2nd ed. Basil Blackwell, 1991.

____ (1979), "SIMSCRIPT", *Encyclopedia of Computer Science and Technology*, Vol. 13, J. Belzer, A.G. Holzman, A. Kent, editors, Marcel Dekker, Inc.

____(1983) *The ER and EAS Formalisms for System Modeling, and the EAS-E Language*, Entity-Relationship Approach to Information Modeling and Analysis, P.P. Chen (ed.) Elsevier Science Publishers B. V. (North-Holland) ER Institute. pp. 29-47.

____ (1991) *Individual versus Institutional Investing*, Financial Services Review, 1(1).

Markowitz, H. M., B. Hausner and H. Karr (1963) SIMSCRIPT: A Simulation Programming Language, Prentice Hall.

Mossin, J. (1968), *Optimal Multiperiod Portfolio Policies*, Journal of Business, 1(2), pp. 215-229.

Samuelson, P.A. (1969), *Lifetime Portfolio Selection by Dynamic Stochastic Programming*, The Review of Economics and Statistics, 51, pp. 239-46.

Exhibit 1 GuidedChoice EAS Table

COM	ENT	ATTR	OWNS	Data/Memb	PREC	COMMENT/Cross References
COM	TYPE			TYPE	(*)	(*) All "Integer"s 32 bit All "Number"s 12,3
ENT	The_SYSTEM					
SET			Sponsors	Sponsor		
SET			Record_keepers	Record_keeper		
SET			GC_AC_frontiers	GC_AC_frontier		
SET			Trust_families	Trust_family		
SET			Security_types	Security_type		
SET			Securities	Security		
SET			Asset_classes	Asset_class		
SET			Partcpnt_auth_memos	Authorization_memo		Authorizastion memos of participants
SET			DB_users	DB_user		Other than participants
SET			Stdd_comp_types	Compensation_type		
SET			Stdd_plan_texts	Plan_text		
SET			Transaction_log_book	Transaction_log_entry		
SET			Event_log_book	Event_log_entry		
SET			Error_log_book	Error_log_entry		
SET			States	State		
ATT		Fed_limit_pretax_dlrs		Number		Currently $10,500
ATT		Fed_limit_pretax_pct		Number		Currently 25%
ATT		Fed_limit_total_dlrs		Number		Currently $30,000
ATT		Fed_limit_total_pct		Number		Currently 25%
SET			Fed_tax_brackets	Inc_tax_bracket		
ENT	State					SYSTEM
ATT		State_id		ID		
ATT		State_name		Text	32	
ATT		State_code		Char	2	
SET			Inc_tax_brackets	Inc_tax_bracket		
ENT	Inc_tax_bracket					SYSTEM, State
ATT		Inc_tax_bracket_id		ID		
ATT		Taxing_government		ID		State_id or -1 for federal govt
ATT		From_income		Number		
ATT		To_income		Number		
ATT		Marginal_rate_pct		Number		

<table>
<tr><th>Kind</th><th>Name</th><th>Type</th><th>Length</th><th>Notes</th></tr>
<tr><td>ENT</td><td>Authorization_memo</td><td></td><td></td><td>SYSTEM, Person</td></tr>
<tr><td>ATT</td><td>Authorization_memo_id</td><td>ID</td><td></td><td>NOT primary key: see next line.</td></tr>
<tr><td>ATT</td><td>Access_SSN</td><td>Char</td><td>11</td><td>Primary key</td></tr>
<tr><td>ATT</td><td>Participant</td><td>Person_id</td><td></td><td>Person</td></tr>
<tr><td>ATT</td><td>Access_pin</td><td>Char</td><td>12</td><td></td></tr>
<tr><td>ATT</td><td>Personal_question</td><td>Text</td><td>128</td><td></td></tr>
<tr><td>ATT</td><td>Secret_answer</td><td>Text</td><td>128</td><td></td></tr>
<tr><td>ATT</td><td>Pin_change_allowed</td><td>Char</td><td>1</td><td></td></tr>
<tr><td>ATT</td><td>Pin_change_date</td><td>Date</td><td></td><td><=Today</td></tr>
</table>

<table>
<tr><th>Kind</th><th>Name</th><th>Type</th><th>Length</th><th>Notes</th></tr>
<tr><td>ENT</td><td>Person</td><td></td><td></td><td>||Person, Authorization_memo,GC_case,Session,</td></tr>
<tr><td>ATT</td><td>Person_id</td><td>ID</td><td></td><td>>Transaction_log_entry, Event_log_entry</td></tr>
<tr><td>ATT</td><td>First_name</td><td>Text</td><td>32</td><td>>Error_log_entry, Expenses_worksheet</td></tr>
<tr><td>ATT</td><td>Last_name</td><td>Text</td><td>32</td><td></td></tr>
<tr><td>ATT</td><td>Middle_initial</td><td>Char</td><td>1</td><td></td></tr>
<tr><td>ATT</td><td>Birth_date</td><td>Date</td><td></td><td><Today</td></tr>
<tr><td>ATT</td><td>Soc_sec_num</td><td>Char</td><td>11</td><td></td></tr>
<tr><td>ATT</td><td>Gender</td><td>Char</td><td>1</td><td>.=MF</td></tr>
<tr><td>ATT</td><td>Contact_info</td><td>Contact_info_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td>Country_of_legal_residence</td><td>Text</td><td>32</td><td></td></tr>
<tr><td>ATT</td><td>Marital_status</td><td>Char</td><td>1</td><td>YN</td></tr>
<tr><td>ATT</td><td>Why_in_DB</td><td>Char</td><td>1</td><td>.=PSB(Partcpnt, Spouse,both)</td></tr>
<tr><td>ATT</td><td>Other_income</td><td>Number</td><td></td><td>Not in Account's Person_comps</td></tr>
<tr><td>ATT</td><td>Total_income</td><td>Number</td><td></td><td></td></tr>
<tr><td>ATT</td><td>Retirement_age</td><td>Integer</td><td></td><td></td></tr>
<tr><td>ATT</td><td>Ret_income_goal_DorP</td><td>Char</td><td>1</td><td>D|P</td></tr>
<tr><td>ATT</td><td>AT_ret_income_goal</td><td>Number</td><td></td><td></td></tr>
<tr><td>ATT</td><td>Spouse</td><td>person_id</td><td></td><td>Person</td></tr>
<tr><td>ATT</td><td>Expenses_worksheet</td><td>Expenses_worksheet_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td>Nr_dependents</td><td>Integer</td><td>2</td><td></td></tr>
<tr><td>SET</td><td>Sign_ons</td><td>Authorization_memo</td><td></td><td>>=1 source doc says this is set?</td></tr>
<tr><td>SET</td><td>Dependents</td><td>Dependant</td><td></td><td>?should this be hooked onto marriage?</td></tr>
<tr><td>SET</td><td>GC_accounts</td><td>Account</td><td></td><td></td></tr>
<tr><td>SET</td><td>Portfolios</td><td>Portfolio</td><td></td><td></td></tr>
<tr><td>SET</td><td>Positions</td><td>Position</td><td></td><td>.= positions in all of Person's portfolios</td></tr>
<tr><td>SET</td><td>Pensions</td><td>Pension</td><td></td><td></td></tr>
<tr><td>SET</td><td>Planned_disbursements</td><td>Planned_disbursement</td><td></td><td></td></tr>
<tr><td>SET</td><td>Non_GC_plans</td><td>Non_GC_plan</td><td></td><td>Plans of Person rprsntng spouse go here</td></tr>
</table>

ENT	Dependent				||Planned_disbursement
ATT		Dependent_id	ID		
ATT		Participant	Person_id		
ATT		Dependent_name	Text	32	
ATT		Dependent_birthdate	Date		>=Today
ATT		Dependent_gender	Char	1	.=M|F
ENT	Planned_disbursement				Person
ATT		Planned_disbursement_id	ID		
ATT		Participant	Person_id		
ATT		Dependent	Dependent		.=0 if not for college
ATT		Disbursement_period	Char		.=M|Q|S|Y
ATT		Disbursement_amt	Number		
ATT		Inflation_adjust_amt	Char	1	YN
ATT		Start_date	Date		
ATT		End_date	Date		>=Start_date
ATT		Disbursement_type	Integer		
ATT		Disbursement_name	Text	128	
ENT	Account				Plan, Person || Pension, Session
ATT		Account_id	ID		>Transaction_log_entry,
ATT		Participant	Person_id		
ATT		Plan	Plan_id		
ATT		Sponsor	Sponsor_id		
ATT		Contact_info	Contact_info_id		
ATT		Tax_state	Text	32	1 of 2 !=""
ATT		Spouse	Spouse_id		
ATT		Employee_status	Enum		1 Active 2 Hardship 3 Terminated
ATT		Eligibility_date	Date		
ATT		Eligibility_match	Char	1	YN
ATT		Eligibility_pension	Char	1	YN
ATT		Years_service	Integer		
ATT		Highly_compensated	Char	1	YN
ATT		Hire_date	Date		<=Today
ATT		GC_advice_accepted	Char	1	YN
ATT		Start_advice_date	Date		
ATT		End_advice_date	Date		
ATT		Eligibility_profit_share	Char	1	YN
ATT		Phone_access	Char	1	YN

<table>
<tr><td>ATT</td><td></td><td>Annual_salary</td><td>Number</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Pretax_earnings</td><td>Number</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Posttax_earnings</td><td>Number</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Last_use</td><td>Date</td><td></td><td><=Today</td></tr>
<tr><td>ATT</td><td></td><td>RK_update_date</td><td>Date</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Accepted_case</td><td>GC_case_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Date_case_accepted</td><td>Date</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Base_case</td><td>GC_case_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Initial_advice</td><td>GC_case_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Modified_advice</td><td>GC_case_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Next_case</td><td>GC_case_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Last_session</td><td>Session_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>OK_rcvd_prime_bnf_NE_spouse</td><td>Char</td><td>1</td><td>YN OK received for prime beneficiary not spouse</td></tr>
<tr><td>SET</td><td></td><td>Beneficiaries</td><td>Beneficiary</td><td></td><td></td></tr>
<tr><td>SET</td><td></td><td>Portfolios</td><td>Portfolio</td><td></td><td></td></tr>
<tr><td>SET</td><td></td><td>Person_Comp_types</td><td>Person_Comp_type</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Contrib_spec_PorD</td><td>Char</td><td>1</td><td>P|D If Plan permits either P (%) or D ($)</td></tr>
<tr><td>SET</td><td></td><td>Current_contribs</td><td>Contrib_instruction</td><td></td><td></td></tr>
<tr><td>SET</td><td></td><td>BT_contrib_allocs</td><td>Contrib_allocation</td><td></td><td></td></tr>
<tr><td>SET</td><td></td><td>AT_contrib_allocs</td><td>Contrib_allocation</td><td></td><td></td></tr>
<tr><td>SET</td><td></td><td>PS_contrib_allocs</td><td>Contrib_allocation</td><td></td><td></td></tr>
<tr><td>SET</td><td></td><td>Archived_cases</td><td>GC_case</td><td></td><td></td></tr>
<tr><td>ENT</td><td>Portfolio</td><td></td><td></td><td></td><td>Person, Account||Contrib_instruction</td></tr>
<tr><td>ATT</td><td></td><td>Portfolio_id</td><td>ID</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Participant</td><td>Person_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Account</td><td>Account_id</td><td></td><td>NULL if not owned by GC_Accouont</td></tr>
<tr><td>ATT</td><td></td><td>Tax_type</td><td>Investment_tax_type_id</td><td></td><td>Enum in EJB; Entity in Administrator</td></tr>
<tr><td>ATT</td><td></td><td>Portfolio _name</td><td>Text</td><td>128</td><td></td></tr>
<tr><td>ATT</td><td></td><td>Accum_AT_contrib</td><td>Number</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Monthly_planned_contrib_dlrs</td><td>Number</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Inflation_adjust_contrib</td><td>Char</td><td>1</td><td>YN</td></tr>
<tr><td>SET</td><td></td><td>Positions</td><td>Position</td><td></td><td></td></tr>
<tr><td>SET</td><td></td><td>AC_exposures</td><td>Exposure</td><td></td><td></td></tr>
<tr><td>ENT</td><td>Position</td><td></td><td></td><td></td><td>Person, Portfolio||Kept_investment</td></tr>
<tr><td>ATT</td><td></td><td>Position_id</td><td>ID</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Security</td><td>Security_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Portfolio</td><td>Portfolio_id</td><td></td><td></td></tr>
<tr><td>ATT</td><td></td><td>Person</td><td>Person_id</td><td></td><td></td></tr>
</table>

ATT		Security_type	Security_ty	32	Must be on security type list
ATT		Tax_type	Enum		
ATT		Quantity	Number		Shares or face value
ATT		Valuation_method	Enum		Mkt price, user price, user total
ATT		User_supplied_price	Number		
ATT		Date_of_user_info	Date		
ATT		Total_value	Number		
ATT		Restricted_for_participant	Char	1	YN Company requirement
ATT		Date_unrestricted	Date		
ATT		Total_cost_basis	Number		
ENT	Person_Comp_type				Person, Compensation type Combo
ATT		Person_Comp_type_id	ID		
ATT		Account	Account_id		
ATT		Compensation_type	Compensation_type_id		
ATT		Compensation_type_name	Text	128	
ATT		Pay_periods_per_year	Integer		
ATT		Amount_per_pay_period	Number		
ENT	Compensation_type				SYSTEM, Plan, Eligible_comp_type ||
ATT		Compensation_type_id	ID		>Person_Comp_type, Contrib_instruction
ATT		Plan	Plan_id		
ATT		Compensation_type_name	Text	32	
ENT	Eligible_comp_type				Plan(2) || PCH_ECT,Contrib_instruction
ATT		Eligible_comp_type_id	ID		>Savings_rate_spec
ATT		Plan	Plan_id		
ATT		Eligible_comp_type_name	Text	32	
SET		Compensation_types	Compensation_type		Overlapping sets

Figure 1

GuidedChoice Plan Client
Retirement Savings Plan Return Distributions
5.5 year returns to 6/30/2013, annualized
Active Participants