

HEC Paris
Grande Ecole
Majeure Finance

Marchés des contrats à terme sur l'élection présidentielle américaine de 2008 et influence des sondages d'opinion*

Jean-Christophe BRETON

Sous la direction d'Evren Örs,
Professeur Assistant au Département Finance, HEC Paris.

Le 24 mai 2009.

* Je tiens à remercier M. Evren ÖRS de m'avoir permis de travailler sur un sujet aussi passionnant. J'adresse également mes remerciements à Mrs Jean-Michel BRETON et Thibault LECLERC pour leur aide avisée dans le travail de collecte et de mise en forme des données.

Pour contacter l'auteur : Jean-Christophe BRETON – 14 Rue du pont rubeau – 91360 Epinay sur Orge.
Tél. : 06.98.29.77.33 – E-mail : jeanchristophe.breton@gmail.com

Marchés des contrats à terme sur l'élection présidentielle américaine de 2008 et influence des sondages d'opinion

Jean-Christophe BRETON

Résumé

Nous étudions dans ce mémoire la capacité prédictive des marchés politiques de l'Université de l'Iowa sur les élections présidentielles américaines dans le cadre du scrutin de 2008. A court, moyen et long termes, le marché offre une prévision précise du résultat du scrutin. Nous observons que, dans certaines conditions, les sondages peuvent néanmoins délivrer de meilleures prévisions de l'issue de l'élection. Nous montrons enfin que les résultats des enquêtes d'opinion publiées au cours de la campagne n'ont pas d'impact direct sur les prix des transactions de ces marchés.

Table des matières

RESUME	1
TABLE DES MATIERES	2
1. INTRODUCTION	4
2. PRESENTATION DE L'IOWA ELECTRONIC MARKETS.....	6
2.1. ASPECTS GENERAUX	6
2.2. FONCTIONNEMENT DU MARCHÉ.....	6
2.2.1. <i>Description des contrats portant sur les élections présidentielles américaines.....</i>	<i>7</i>
2.2.2. <i>Achat et vente des contrats.....</i>	<i>8</i>
2.2.3. <i>Contraintes imposées aux investisseurs</i>	<i>8</i>
3. PERFORMANCE HISTORIQUE DES MARCHES DE L'IEM.....	10
3.1. PRECISION DES PREVISIONS DEDUITES DES PRIX DE MARCHÉ	10
3.1.1. <i>Fondements théoriques</i>	<i>10</i>
3.1.2. <i>Performance absolue.....</i>	<i>11</i>
3.1.3. <i>Performance relative.....</i>	<i>11</i>
3.2. COMMENT EXPLIQUER UNE TELLE PERFORMANCE ?	12
3.3. BIAIS ET ANOMALIES DU COMPORTEMENT DES INVESTISSEURS	13
3.4. EXPLICATIONS AVANCEES ET ELEMENTS DE RECHERCHE A POURSUIVRE	14
4. PRESENTATION DES DONNEES.....	15
4.1. SELECTION DES DONNEES	15
4.1.1. <i>Données de marché.....</i>	<i>15</i>
4.1.2. <i>Des sondages nationaux hétérogènes.....</i>	<i>17</i>
4.1.3. <i>Un large échantillon de sondages locaux</i>	<i>18</i>
4.2. ANALYSE DE L'ECHANTILLON	19
4.2.1. <i>Données chiffrées.....</i>	<i>19</i>
4.2.2. <i>Volatilité et versatilité des sondages nationaux.....</i>	<i>20</i>
4.2.3. <i>Une inégale répartition des sondages locaux</i>	<i>22</i>
5. RESULTATS DE LA RECHERCHE.....	23
5.1. EN 2008, LES MARCHES ONT-ILS PREVU LE RESULTAT DU SCRUTIN AVEC PRECISION ?.....	24
5.1.1. <i>Qualité de la prévision à court, moyen et long termes.....</i>	<i>24</i>
5.1.2. <i>Éléments d'explications et conclusions.....</i>	<i>27</i>
5.2. LE POUVOIR DE PREDICTION DU MARCHÉ EST-IL SUPERIEUR A CELUI DES SONDAGES D'OPINION ?.....	28
5.2.1. <i>Méthodologie.....</i>	<i>28</i>

5.2.2.	<i>Quand le marché surpasse les sondages</i>	30
5.2.3.	<i>L'année 2008 n'est pas propice au marché</i>	31
5.2.4.	<i>Conclusion</i>	33
5.3.	LES SONDAGES INFLUENCENT-ILS LES PRIX DE MARCHÉ ?	33
5.3.1.	<i>Le modèle de Forsythe, Nelson, Neumann et Wright (1992)</i>	34
5.3.1.1.	<i>Présentation du modèle</i>	34
5.3.1.2.	<i>Le modèle est invalidé</i>	35
5.3.1.3.	<i>Les sondages sont-ils pour autant inutiles ?</i>	36
5.3.2.	<i>Pouvons-nous construire un autre modèle ?</i>	37
5.3.2.1.	<i>Hypothèses</i>	37
5.3.2.2.	<i>Les variables évidentes ne donnent pas de résultat</i>	38
5.3.2.3.	<i>Impact de l'économie sur le prix de marché</i>	39
5.3.2.4.	<i>L'impact des sondages n'est pas mesurable</i>	40
5.4.	CERTAINS ÉTATS ONT-ILS PLUS D'INFLUENCE QUE D'AUTRES SUR LES PRIX DU MARCHÉ ?	41
5.4.1.	<i>Les Swing States : la clé du prix de marché ?</i>	41
5.4.2.	<i>Des résultats contraires aux hypothèses sur le contrat démocrate</i>	42
5.4.3.	<i>Vérification des résultats par analyse du contrat REP08_WTA</i>	44
5.4.4.	<i>Conclusion</i>	45
6.	CONCLUSION	46
	BIBLIOGRAPHIE	48
	ANNEXES	51

1. Introduction

Depuis l'élection présidentielle de 2002 et le fameux premier tour du 21 avril, les électeurs français ont appris à se méfier des prédictions des instituts de sondages. Beaucoup d'entre eux considèrent que les résultats des sondages ne sont pas fiables et que les entreprises qui publient ces résultats, toutes scientifiques que leurs méthodes puissent être, sont incapables de prévoir l'issue du scrutin, même à court terme. D'ailleurs, comment d'autres pourraient leur annoncer le vainqueur de l'élection alors qu'eux-mêmes ne sont encore pas parvenus à faire un choix ? Il existe pourtant, aux Etats-Unis, un organisme qui parvient à prévoir le résultat de chaque élection présidentielle avec une précision insolente, et ce depuis bientôt vingt ans. Cet organisme, l'Université de l'Iowa, n'utilise aucune méthode statistique d'échantillonnage de la population pour réussir cette prouesse. Il donne simplement la liberté à des investisseurs du monde entier d'acheter et de vendre des contrats à terme, dont les formules de liquidation dépendent de l'issue du scrutin, sur un marché spécifiquement dédié.

Lors de la première ouverture de ce marché, pour l'élection présidentielle de 1988, la pertinence de cet outil de prédiction a surpris même les créateurs du projet. Depuis, et à de multiples reprises, la performance du marché n'a cessé d'étonner les chercheurs, et reléguer en seconde classe les prévisions des instituts de sondages. Ceux-ci n'ont réussi que très rarement à surperformer le marché, si l'on en croit la plupart des travaux de recherche. Pourtant, la suprématie du marché sur les sondages fait encore débat : les défenseurs de ces derniers accusent leurs opposants d'utiliser des hypothèses trop restrictives qui favorisent le marché. Nous nous proposons d'alimenter le débat en étudiant la qualité des prévisions issues du marché et des résultats des sondages, dans le cadre de l'élection présidentielle américaine de 2008. Dans un second temps, nous avons pour objectif de mesurer l'impact des résultats des sondages sur les prix relevés sur les marchés de l'Université de l'Iowa.

Après avoir brièvement décrit le fonctionnement des marchés électroniques de l'Université de l'Iowa, nous présentons les principales caractéristiques des contrats à terme sur l'élection de 2008 et mettons en lumière les spécificités de ces produits qui leur permettent de constituer de bons outils de prédiction du résultat de l'élection. Nous analysons par la suite les principales conclusions des recherches qui ont été menées sur les marchés électroniques de l'Université de l'Iowa, et en particulier ceux traitant de la performance de ces marchés et des raisons qui expliquent la précision des prévisions qui en sont dérivées.

Nous détaillons ensuite la qualité de la précision des prévisions du marché à court terme, à moyen terme et à long terme, de façon absolue (par rapport au résultat de l'élection) et de façon

relative (par rapport aux prévisions établies par les instituts de sondage). Nous montrons que le marché est effectivement plus précis que les sondages lorsque l'on compare les résultats de chaque sondage au marché. En revanche, si l'on utilise la moyenne des résultats publiés chaque jour, le marché s'avère moins précis que l'agrégat de sondages, dans le dernier mois de la campagne électorale.

Malgré cette performance des sondages dans la dernière ligne droite de la campagne, nous observons que nous ne pouvons pas établir de lien entre les résultats des sondages, sous quelque forme que ce soit, et les prix de marché. Nous remarquons en revanche que l'utilisation d'indicateurs économiques dans le modèle pourrait permettre de reproduire, en partie, les variations de prix relevées sur le marché. Nous notons également que le prix des contrats semble suivre une marche aléatoire.

Les marchés électroniques sur les élections présidentielles de l'Université de l'Iowa sont une des premières manifestations des marchés prédictifs, dont le but principal n'est pas d'allouer les ressources de façon optimale comme dans un marché traditionnel, mais de prévoir un événement futur. Le champ d'application de ces marchés dépasse de loin celui de la politique. Schonfeld (2006) rapporte que Hewlett-Packard a permis à ses managers d'échanger des contrats sur un marché interne dans le but de prévoir le prix à terme d'un des composants essentiels de ses produits, les processeurs. Au cours des années 1990, la même entreprise testait déjà avec succès un marché destiné à prévoir les ventes de PC de la marque. Dans les marchés internes créés par Google, les employés échangent des contrats à terme dans le but de prévoir le nombre d'utilisateurs des nouvelles applications développées par le groupe. Nul besoin d'étendre la liste des exemples pour comprendre que les marchés prédictifs sont destinés à être largement utilisés à l'avenir. Il est donc essentiel de comprendre leur fonctionnement aujourd'hui, pour mieux utiliser tout leur potentiel demain. C'est l'objectif que nous nous proposons de partager, avec notre lecteur, tout au long de ce mémoire de recherche.

2. Présentation de l'Iowa Electronic Markets

2.1. Aspects généraux

Créés pour la première fois en 1988, les plateformes électroniques d'échange de l'Université de l'Iowa permettent d'échanger des contrats à terme (*futures*) dont le règlement est déterminé par la réalisation *in fine* d'événements, qu'ils soient politiques (résultats des élections), économiques (valeur de l'action Microsoft) ou culturels (contrats futures indexés sur les performances de certains films au Box-Office américain). Les marchés de l'Iowa Electronic Markets (IEM) ont été conçus à l'origine par les membres du corps professoral et de l'administration du Henry B. Tippie College of Business de l'Université de l'Iowa à des fins de recherche et d'enseignement. Encore aujourd'hui, les traders qui n'appartiennent pas à la communauté de l'Université de l'Iowa ne sont pas autorisés à intervenir sur l'IEM, à l'exception des marchés sur les élections politiques.

Les marchés de l'IEM¹ permettent aux investisseurs de négocier des contrats à terme, en temps réel et en investissant leurs ressources financières personnelles. Les investisseurs accèdent au marché en temps continu, via Internet, à partir de n'importe quel point du globe. Chaque acteur du marché dûment inscrit possède un compte de dépôt sur lequel il peut placer un montant minimum de 5\$ et maximum de 500\$, ce compte pouvant être alimenté en plusieurs fois. Des frais de 5\$ sont prélevés à l'ouverture du compte et tous les six mois sur les comptes inactifs.

L'investisseur inscrit auprès de l'IEM a accès aux informations de marché en temps continu : prix du contrat déterminé lors de la dernière transaction, meilleurs prix à la vente (*bid*) et à l'achat (*ask*) du carnet d'ordres. En revanche, l'IEM ne lui donne pas la possibilité de mesurer la profondeur du carnet d'ordres : l'investisseur ne sait ni combien de contrats sont proposés aux meilleurs *bid* et *ask*, ni quels sont les ordres à cours limités placés au-delà des meilleures cotations.

2.2. Fonctionnement du marché

Nous décrivons tout d'abord les caractéristiques des contrats échangés sur l'IEM ainsi que les règles de trading en vigueur sur le marché et leurs éventuelles implications pour les investisseurs.

¹ Bien que l'ensemble des contrats de l'IEM représente un potentiel de recherche considérable, nous limitons ici notre étude aux marchés des élections politiques et toute référence ultérieure à l'IEM concernera exclusivement ces marchés.

Le lecteur trouvera en annexe une capture d'écran d'un marché de l'IEM tel qu'il apparaît à tout investisseur voulant prendre position sur l'IEM.

2.2.1. Description des contrats portant sur les élections présidentielles américaines

Les deux marchés de l'IEM destinés aux élections présidentielles américaines, le *Vote Share Market* (marché où les contrats ont une valeur qui est proportionnelle au pourcentage des votes que le candidat recevra lors de l'élection) et le *Winner-Take-All Market* (marché où seul le contrat du candidat vainqueur de l'élection aura une valeur de liquidation strictement positive), ont permis l'échange de contrats à terme sur une période d'un peu plus de deux ans, entre le 1^{er} juin 2006 et la publication officielle des résultats, soit le 7 novembre 2008. Les contrats portant sur les élections présidentielles américaines sont échangés sur deux marchés et distingués suivant la formule de liquidation utilisée à l'échéance. En 2008, contrairement aux précédentes éditions, seuls deux contrats étaient proposés sur chaque marché.

Sur le *Vote Share Market*, les investisseurs pouvaient prendre position sur le contrat UDEM08_VS dont la formule de liquidation est indexée sur le pourcentage des suffrages exprimés pour le candidat du parti démocrate, et sur le contrat UREP08_VS dont la formule de liquidation est indexée sur le pourcentage des suffrages exprimés pour le candidat républicain. Pour chacun de ces deux contrats, les formules de liquidation à l'échéance sont les suivantes :

$$UDEM08_VS = \frac{DEM\%}{DEM\% + REP\%} \cdot 1\$ \quad (1)$$

$$UREP08_VS = \frac{REP\%}{REP\% + DEM\%} \cdot 1\$ \quad (2)$$

où DEM% et REP% représentent respectivement les pourcentages des suffrages exprimés lors de l'élection présidentielle du 4 novembre 2008. Le pourcentage des suffrages exprimés pour un candidat est ramené à la somme des pourcentages exprimés pour chacun des deux principaux candidats de sorte qu'un portefeuille contenant une unité de chaque contrat ait une valeur d'exactly 1\$ à la liquidation. Par construction, la somme des prix des deux contrats doit donc être constamment proche de 1\$.

Sur le *Winner-Take-All Market*, les investisseurs pouvaient prendre position sur le contrat DEM08_WTA représentant le candidat démocrate, et sur le contrat REP08_WTA représentant le candidat républicain. Le détenteur d'un de ces contrats reçoit 1\$ à la date de liquidation si le candidat associé est déclaré vainqueur de l'élection présidentielle, 0\$ dans les autres cas. Les formules de liquidation sont établies de sorte qu'un investisseur qui posséderait une unité de chaque contrat à l'échéance reçoive 1\$ de façon certaine, quelle que soit l'issue du vote.

2.2.2. Achat et vente des contrats

Les contrats échangés sur le marché peuvent être achetés ou vendus à l'IEM sous forme de portefeuilles, appelés *unit portfolios*, comportant une unité de chaque contrat listé sur un marché. Par exemple, sur le *Vote Share Market*, tout investisseur peut acheter ou vendre un portefeuille composé d'un contrat UDEM08_VS et d'un contrat UREP08_VS à l'IEM pour une valeur de 1\$ (d'où le nom de *unit portfolio*, ou portefeuille unité). Cette règle a trois implications essentielles pour le fonctionnement du marché. Tout d'abord, elle garantit que l'IEM, qui gère les marchés, ne peut espérer aucun gain ou perte financière lié à l'achat ou à la vente des contrats. Ensuite, elle implique qu'il y ait à tout moment le même nombre de contrats « démocrates » et « républicains » en circulation sur le marché, le prix du contrat n'étant pas contraint par la quantité de contrats disponibles (on peut toujours créer un nouveau contrat en échangeant 1\$ de son compte de dépôt contre un *unit portfolio*). Enfin, le fait que tout portefeuille contenant une unité de chaque contrat puisse être acheté ou vendu à tout moment à un prix fixe et certain de 1\$ implique que le taux d'intérêt sans risque sur l'IEM est nécessairement nul.

A l'instar des marchés financiers du monde entier, un investisseur peut acheter ou vendre un contrat en envoyant au marché un ordre au mieux (*market order*) ou un ordre à cours limité (*limit order*). L'ordre à cours limité (constitué d'un prix, d'une quantité et d'une date de validité) est placé dans le carnet d'ordres s'il ne peut s'exécuter instantanément contre un ordre opposé du carnet d'ordres. L'ordre au mieux est exécuté uniquement aux cotations qui prévalent au moment de la transmission de l'ordre : si la quantité d'achat de contrats (resp. de vente) stipulée dans l'ordre au mieux est supérieure à la quantité proposée au meilleur *ask* (resp. *bid*) alors la partie de l'ordre au mieux non exécutée est annulée par l'IEM. Un investisseur qui souhaiterait acquérir une quantité déterminée de contrats à une date donnée préférera donc placer un ordre à cours limité à un meilleur prix que celui indiqué par la fourchette de prix au moment de passer l'ordre.

2.2.3. Contraintes imposées aux investisseurs

Bien qu'exerçant dans les faits les fonctions d'un marché de contrats à terme, l'IEM reste aujourd'hui encore en dehors du système de régulation américain². Cette particularité tient

² L'IEM indique sur son site Internet : "The IEM is not regulated by, nor are its operators registered with, the Commodity Futures Trading Commission or any other regulatory authority. The IEM has received two no-action letters from the Division of Trading and Markets of the Commodity Futures Trading Commission. Without explicitly asserting jurisdiction over the IEM or any of its submarkets, these letters, dated February 5, 1992, and June 18, 1993, extended no-action relief to the IEM's Political and Economic Indicator Markets." Traduction : *L'IEM n'est pas régulé, ni inscrit, auprès de la Commodity Futures Trading Commission ou de tout autre autorité de régulation. L'IEM a reçu*

notamment au fait que le montant total des investissements de chaque individu ne peut dépasser 500\$.

Tous les achats doivent être réalisés au comptant et les ventes à découvert sont interdites. Le système évalue pour chaque unité échangée la faisabilité de l'échange en fonction des contraintes imposées par le marché aux investisseurs. Cette procédure assure que les règles de fonctionnement du marché sont respectées. Un investisseur peut par exemple placer un ordre de vente pour un nombre de contrat supérieur à celui qu'il possède effectivement en portefeuille ; néanmoins, le système n'exécutera la vente que du nombre de contrat qu'il possède réellement et annulera le reste de l'ordre non exécuté. Bien que les ventes à découvert soient proscrites, tout investisseur a la possibilité de répliquer de façon synthétique une vente à découvert en achetant des *unit portfolios* auprès de l'IEM et en revendant directement sur le marché les titres qu'il souhaitait vendre à découvert. D'autre part, dans un marché constitué de seulement deux actifs, tels que les marchés des élections présidentielles américaines en 2008, être vendeur d'un contrat revient à adopter la position acheteur sur le contrat « complémentaire ». En effet, un investisseur souhaitera vendre un contrat lorsqu'il estime que celui-ci est surévalué, ce qui revient à dire que le second contrat est nécessairement sous-évalué et que, suivant son raisonnement, il devrait être à l'achat sur ce contrat.

Lorsque l'on découvre le fonctionnement de l'IEM, on pourrait penser que les prix des contrats sont facilement manipulables par les acteurs du marché. A priori, la limitation du montant total d'investissement à 500\$ doit empêcher tout individu d'avoir un impact durable sur les prix, à moins de multiplier les comptes ouverts pour un même individu. Par ailleurs, il est impossible pour un même investisseur d'avoir les positions acheteur et vendeur lors d'un échange : lorsque le marché est peu liquide, cette règle empêche un investisseur de réaliser des échanges à des cours artificiellement élevés sans encourir de perte financière. Enfin, si le marché est efficient, la présence d'arbitrageurs doit empêcher une déviation trop importante des prix de la valeur fondamentale. Nous reviendrons sur l'efficacité de l'IEM et la question de la manipulation des prix dans la section suivante.

deux lettres de la Division du Trading et des Marchés de la Commodity Futures Trading Commission indiquant qu'elle n'intenterait pas d'action réglementaire à l'égard de l'IEM. Sans affirmer explicitement la compétence de cette autorité sur l'IEM ou ses marchés, ces lettres, datées du 5 février 1992 et du 18 juin 1993, ont étendu l'exception aux marchés politiques et économiques de l'IEM.

<<http://www.biz.uiowa.edu/iem/faq.html>>

3. Performance historique des marchés de l'IEM

A l'origine, l'IEM a été créé à des fins pédagogiques et de recherche. L'objectif était à la fois de pouvoir étudier le fonctionnement des marchés prédictifs et le comportement des investisseurs actifs sur ces marchés. Dans cette section, nous analysons les travaux de recherche portant sur les précédentes éditions des marchés de l'IEM sur les élections présidentielles (de 1988 à 2004). Nous verrons tout d'abord les arguments théoriques qui justifient la performance de ces marchés, à la fois en terme absolu (lorsque l'on compare le prix de marché au résultat de l'élection) et en terme relatif (par rapport à d'autres outils prédictifs, et notamment les sondages d'opinion). Nous présentons ensuite le paradoxe posé par les résultats de cette analyse de performance, avant de détailler les explications qui ont été jusqu'ici proposées et les éléments qui restent encore à étudier.

3.1. Précision des prévisions déduites des prix de marché

3.1.1. Fondements théoriques

Les contrats et les formules de liquidation associées sont déterminés de sorte que le prix de marché puisse être considéré comme une prévision du résultat final de l'élection. Berg, Nelson et Rietz (1993) rappellent que cette conclusion est dérivée des modèles du CAPM et de l'Arbitrage Pricing Theory. Dans le cadre de ces théories, la valeur du contrat est égale à l'espérance du prix à l'échéance, actualisée au taux de rentabilité exigé. Or, comme le taux de rentabilité du portefeuille de marché et le taux sans risque sont nuls, toute prime de risque est nulle. Il en résulte, d'après le CAPM, que tout actif doit avoir une espérance de rentabilité nulle. Le prix du contrat à une date t est donc réduit à l'espérance du prix de ce contrat à une date future. En particulier, le prix actuel représente donc l'espérance du prix du contrat à l'échéance, et par conséquent, l'espérance du résultat de l'élection. La théorie soutient donc la performance absolue des prix de marché comme éléments de prévision des résultats de l'élection présidentielle.

La performance relative mesure l'écart de performance entre les prix de marché et d'autres éléments de prédiction, et notamment les plus utilisés d'entre tous : les sondages d'opinion. Kou et Sobel (2001) montrent que, sous certaines hypothèses, les prix de marché conduisent à des erreurs de prévision inférieures aux erreurs de prévision produites par les sondages d'opinion. Les auteurs supposent en particulier que les sommes investies sur l'IEM ne font pas partie de la stratégie d'investissement des individus et, par conséquent, n'ont pas à être arbitrées en fonction des performances des investissements dans les autres classes d'actifs. Ils

soulignent que la limite maximale des 500\$ sur l'PIEM joue un rôle fondamental dans la justification de cette hypothèse.

3.1.2. Performance absolue

A l'exception de l'élection présidentielle de 2000³, le prix de marché s'est toujours révélé une excellente prédiction du résultat du scrutin. Selon Berg, Forsythe, Nelson et Rietz (2000), les erreurs absolues moyennes des prévisions de l'ensemble des marchés de l'PIEM se sont établies à 1,37% pour les élections présidentielles américaines, 3,43% pour les autres élections américaines et 2,12% pour les marchés prévoyant certaines élections en dehors des Etats-Unis (Europe et Canada principalement). Ils montrent que les prix de marché sont de bons éléments de prédiction à court terme (lorsque l'on considère les prix de marché la veille ou la semaine précédent le scrutin) du résultat de l'élection. Berg, Nelson et Rietz (2003) ajoutent que les prix de marché semblent prévoir l'issue du scrutin avec une certaine précision plusieurs mois à l'avance.

3.1.3. Performance relative

Entre 1988 et 2004, Berg, Nelson et Rietz (2008) relèvent qu'en moyenne les sondages d'opinion ont produit une erreur d'estimation du résultat de l'élection présidentielle de 1,62% alors que le marché est parvenu à prédire l'issue du scrutin avec une erreur moyenne de 1,33% (si l'on considère les cotations du marché la veille de l'élection) et de 1,20% (par une moyenne des cours de la semaine précédent l'élection, pondérée par les volumes échangés). Les instituts de sondage s'opposent à une telle comparaison, soutenant que leurs sondages mesurent, au travers de la question « si l'élection présidentielle avait lieu aujourd'hui... ? », les préférences actuelles des électeurs et non celles qui s'exprimeront lors du scrutin. Toutefois, nul ne peut nier que les sondages sont interprétés comme tel. D'autre part, les instituts de sondage avaient auparavant pour habitude de demander directement aux sondés leurs intentions de vote pour le jour de l'élection. Gelman et King (1993) soulignent que le changement de méthode n'a eu d'autre conséquence que d'augmenter la proportion des indécis et que l'on peut valablement supposer que l'on obtiendrait les mêmes résultats en interrogeant les électeurs sur leurs intentions de vote pour le jour du scrutin. Sur des périodes de temps plus longues, Berg, Nelson et Rietz (2003) montrent que le marché est capable de prédire l'issue du scrutin de façon plus précise que les sondages, et ce plusieurs mois à l'avance.

³ En 2000, le *Winner-Take-All Market* prévoyait une victoire du candidat républicain alors que les citoyens américains avaient choisi majoritairement le candidat démocrate. La Cour Suprême américaine a toutefois donné raison aux investisseurs cette année-là.

3.2. Comment expliquer une telle performance ?

Une explication de la meilleure performance des sondages pourrait provenir de la qualité des acteurs intervenant sur le marché. Selon cette hypothèse, l'PIEM aurait réussi la prouesse de réunir « naturellement » un échantillon représentatif de la population américaine. La meilleure performance du marché ne proviendrait que d'une erreur d'estimation plus faible due à l'échantillon. Cette hypothèse est infondée selon Berg et Rietz (2006, pp. 145-146) qui établissent que les investisseurs de l'PIEM ont tendance à être surtout des hommes, plus riches et plus diplômés que l'Américain moyen. D'autre part ; lors de l'édition 1988, seuls les membres de l'Université de l'Iowa étaient autorisés à participer au marché de l'PIEM. Les prévisions de ce marché se sont révélées très précises, alors que l'Iowa représente moins de 1% du corps électoral américain.

Levmore (2003) suggère que l'PIEM aurait beaucoup moins de réussite à prévoir le résultat de l'élection si les sondages n'existaient pas. Il soutient que les acteurs du marché parviennent à des prévisions plus justes en interprétant les sondages et en incorporant ces informations dans le marché. En disposant d'informations supplémentaires (les résultats des sondages) que les sondés n'ont pas à disposition lorsqu'ils répondent aux enquêtes, les investisseurs auraient un avantage lorsqu'il s'agit d'évaluer la valeur fondamentale des contrats. On peut d'ailleurs facilement imaginer que certains investisseurs utilisent principalement les sondages pour prendre leurs décisions d'investissement. Forsythe, Nelson, Neumann et Wright (1992) ne trouvent aucune relation de causalité entre les prix de marché et les résultats des sondages d'opinion et suggèrent que ces derniers sont intégrés par le marché bien avant leur publication.

Après les travaux de Forsythe, Nelson, Neumann et Wright en 1992, la question de la corrélation entre les prix de marché et les sondages d'opinion n'a pas donné lieu à de nouvelles recherches et un consensus semblait s'être dégagé autour des conclusions de leurs travaux. Pourtant, Kou et Sobel (2001) remettent en cause cette vision en réfutant la validité des tests statistiques utilisés par les quatre chercheurs et le caractère restrictif des hypothèses formulées. Ils montrent au contraire que lorsque les sondages et le marché fonctionnent conformément à la théorie, les prévisions du marché ne sont meilleures que les sondages seulement si les investisseurs ont accès aux résultats des sondages (ce qui est vrai dans la pratique). Erikson et Wlezien (2008) affirment que lorsque les sondages d'opinion sont correctement interprétés, ils donnent des résultats beaucoup plus précis que les prix de marché. A cette fin, ils suggèrent que les résultats des sondages apparaissent systématiquement biaisés, à chaque élection, et qu'en tenant compte de ce biais historique, il est possible de corriger les estimations des sondages afin de parvenir à une prévision plus fiable. Nous remarquons cependant que les instituts de sondage

incluent déjà des éléments correctifs qui visent à limiter les biais observés lors des précédents scrutins. En appliquant de multiples altérations aux données initiales, peut-on encore estimer que l'estimation obtenue est le résultat d'un sondage ou devient-elle le fruit d'un modèle agrégeant l'information ? Gelman et King (1993) proposent d'ailleurs de prévoir l'issue du scrutin présidentiel de 1988 à l'aide d'un modèle prédictif, avec un certain succès. Lockerbie (2004 et 2005) s'essaye également au difficile exercice de prévision du résultat du scrutin de 2004 à l'aide d'un modèle basé sur l'anticipation de l'état de l'économie américaine et la lassitude de l'opinion vis-à-vis du parti au pouvoir, sans succès. Quoi qu'il en soit, la performance relative du marché par rapport aux autres moyens de prédiction est une question qui n'est toujours pas résolue. Il nous apparaît donc intéressant d'étudier ce point sur le dernier scrutin à disposition, les élections présidentielles de 2008. A notre connaissance, cette recherche n'a pas encore été conduite.

3.3. Biais et anomalies du comportement des investisseurs

La performance des prévisions du marché surprend d'autant plus que les comportements des investisseurs sont parfois irrationnels. Forsythe, Nelson, Neumann et Wright (1992) et Forsythe, Rietz et Ross (1998) identifient deux biais principaux chez les investisseurs de l'IEM : (1) la tendance des investisseurs à croire que leurs opinions personnelles sont représentatives de la population dans son ensemble (« *false-consensus effect* ») et (2) l'influence des opinions personnelles de l'investisseur sur la façon d'interpréter les informations relatives au scrutin (« *assimilation-contrast effect* »). Consécutivement, Oliven et Rietz (2004) montrent que ces comportements irrationnels conduisent les investisseurs à perdre systématiquement une partie de leur richesse lorsqu'ils échangent sur le marché. Dès lors, comment expliquer que le marché parvienne à prévoir avec succès le résultat d'une élection ? Pour répondre à cette question, Oliven et Rietz (2004) poursuivent leur analyse et scindent l'ensemble des investisseurs en deux groupes : (1) les investisseurs qui placent principalement des ordres au mieux lorsqu'ils échangent des contrats et que les auteurs appellent « preneurs de prix » (« *price takers* ») et (2) les investisseurs qui placent des ordres à cours limité à des prix proches de la fourchette de prix et que les auteurs appellent « faiseurs de prix » (« *market makers* »). Ils montrent que les *price takers* sont davantage susceptibles d'exhiber des comportements irrationnels et d'encourir des pertes financières lorsqu'ils interviennent sur le marché, alors que les actions des *market makers* sont beaucoup plus rationnelles et, en fin de compte, produisent des stratégies avec un résultat financier légèrement positif. Ils concluent qu'en profitant des erreurs des preneurs de prix, les faiseurs de prix ont une incitation à rechercher l'information et rendre le marché plus efficient.

3.4. Explications avancées et éléments de recherche à poursuivre

Outre le fait que les contrats échangés sur le marché soient spécifiquement conçus pour lier le prix de marché à l'issue du scrutin, Berg, Nelson et Rietz (2003) affirment que le marché tire sa précision de trois raisons fondamentales : (1) les caractéristiques du marché forcent les investisseurs à se pencher spécifiquement sur le thème de l'élection présidentielle, (2) les investisseurs sont intéressés au processus de découverte du prix car ils investissent leurs propres économies et (3) le marché agrège les diverses informations des investisseurs de façon dynamique et potentiellement efficiente. Berg, Forsythe et Rietz (1997) montrent que trois facteurs contribuent à la précision des marchés de l'IEM : le nombre de types de contrats en circulation dans un marché (correspondant pratiquement au nombre de candidats sur lesquels il est possible de prendre position), le volume de contrats échangés (mesuré en dollars américains) avant l'élection et le déséquilibre du carnet d'ordre la veille de l'élection (différence entre les volumes pondérés par les prix des ordres d'achat et de vente à cours limité non exécutés la veille de l'élection). Plus cette différence et le nombre de contrats proposés sont élevés, plus l'erreur de prévision est grande. Inversement, plus le volume échangé est élevé, plus l'erreur de prévision est limitée. Ces résultats rejoignent les observations d'Oliven et Rietz (2004) qui relèvent que l'augmentation du volume échangé sur le marché s'accompagne d'un accroissement des comportements biaisés de la part des *price takers*, fournissant ainsi davantage d'incitations aux *market makers* et favorisant le processus de découverte du prix. Kou et Sobel (2001) ajoutent que le marché se révèle capable de prédire le résultat des élections présidentielles notamment parce que les investisseurs actifs sur le marché n'ont pas d'incitation à utiliser le marché dans un autre but que celui de prédire le résultat du scrutin. Si le marché devait servir comme moyen de couvrir d'autres investissements ou de se protéger contre des mouvements de prix sur les marchés financiers, la précision de l'IEM s'avèrerait bien médiocre selon ces deux auteurs.

A la lumière de l'ensemble des recherches déjà menées sur l'IEM, il nous semble intéressant de revenir sur le lien entre les prix de marché et les résultats des sondages d'opinion, sous l'angle de l'élection présidentielle américaine de 2008. En particulier, l'influence des résultats des sondages d'opinion sur les prix de marché reste encore un élément de controverse. Les performances absolues et relatives des prix de marché doivent être comparées avec une base de sondages suffisamment large et répartie au cours du temps pour pouvoir mettre en lumière une corrélation. Si, en théorie, l'efficience des marchés suppose qu'un investisseur ne peut pas profiter d'une information publique (telle que les résultats des sondages), une information publique telle que le résultat d'un sondage doit nécessairement être prise en compte par le marché. Il nous

apparaît essentiel de revenir sur cette question et de l'approfondir. En effet, un certain nombre de recherche ont été menées au niveau fédéral mais, à notre connaissance, aucune recherche n'a été menée concernant l'influence des sondages réalisés dans les Etats fédérés. Dans le cas particulier des Etats-Unis, les journalistes et les observateurs politiques soulignent souvent l'influence de certains Etats clés, les *Swing States*, pour l'issue du vote. Nous tâcherons d'étudier l'influence des résultats des sondages d'opinion conduits dans ces Etats sur les prix du marché.

4. Présentation des données

Nous proposons tout d'abord de décrire notre procédure de collecte des données et de sélection des éléments à incorporer dans notre panel. Nous pensons en effet que ce processus garantit à la fois la qualité des données brutes et des résultats que nous présenterons par la suite. Consécutivement, nous menons une analyse détaillée de l'échantillon afin d'en illustrer les qualités mais aussi les limites et les problèmes posés par l'hétérogénéité des données.

4.1. Sélection des données

Afin de remplir les objectifs de recherche définis à la section précédente, nous avons réuni trois types de données : les données de marché, les résultats des sondages d'opinion établis au niveau fédéral (que nous désignerons ci-après comme les « sondages nationaux ») et les résultats des sondages d'opinion établis au niveau des états fédérés (ci-après, les « sondages locaux »). Nous détaillons dans les sous-sections suivantes le processus de sélection et d'obtention des données, ainsi que la logique qui sous-tend leur mise en valeur.

4.1.1. Données de marché

L'IEM offre, en accès plus ou moins libre, un certain nombre d'informations concernant ses marchés et notamment des données quotidiennes historiques relatives aux marchés politiques. Ces données sont accessibles directement via le site Internet de l'IEM⁴. Pour chaque marché et à chaque nouvelle élection, l'IEM propose généralement un relevé quotidien de données regroupant le nombre total de contrats échangés au cours de la journée, la valeur totale de ces échanges (en \$), les cours les plus bas et les plus élevés de la journée, le prix moyen auquel les échanges ont eu lieu et le prix de clôture. Bien que le format de présentation des données historiques soit relativement flexible, il respecte généralement le schéma que nous venons de

⁴ Le site Internet de l'IEM est accessible à l'adresse : <www.biz.uiowa.edu/iem>

décrire, ce qui nous permet de conserver une certaine cohérence entre les données de l'édition 2008 et les données des anciennes élections présidentielles américaines.

Les prix de marché que nous utilisons dans le cadre de notre recherche sont les prix quotidiens relevés à la clôture par l'IEM, et tels que définis par le prospectus afférant à chaque contrat échangé sur le marché. En effet, les marchés de l'IEM opèrent sans discontinuer, vingt-quatre heures sur vingt-quatre et sept jours sur sept. L'IEM définit le prix de clôture d'un contrat comme le prix relevé lors de la dernière transaction, avant minuit. Cette heure a été choisie par l'IEM car elle correspond au moment où les ordres à cours limité non exécutés et à date de validité échue sont retirés du carnet d'ordres. Lorsque aucune transaction n'a eu lieu au cours de la journée, le prix de clôture du jour précédent est reconduit comme nouveau prix de clôture.

Cette définition du prix de clôture n'est pas exempte d'imperfections. Tout d'abord, dans un marché peu liquide, elle peut conduire à une représentation artificielle des prix. En effet, si une transaction a eu lieu au cours d'une journée à un prix donné, et s'il n'y a pas de transaction au cours des trois ou quatre jours suivants, ce prix sera considéré comme le prix de clôture du marché, même s'il ne reflète pas l'état du marché tel que défini par les ordres à cours limité enregistrés dans le carnet d'ordres. Bien que ce cas soit relativement rare, nous avons relevé quelques exemples de ce problème dans la série des prix du *Vote Share Market*, notamment entre le 1^{er} mai 2008 et le 4 mai 2008. Le 1^{er} mai 2008, une transaction a eu lieu avant la clôture sur le contrat UREP08_VS à un prix que nous estimons artificiellement élevé de 0,74\$ (qui doit être interprété comme un pari de l'acheteur du contrat sur une victoire du candidat républicain aux élections présidentielles, avec au moins 74% des suffrages exprimés en sa faveur⁵). Consécutivement, aucune transaction n'a eu lieu entre le 2 et le 4 mai 2008. L'IEM considère donc que, pour ces trois journées, le prix de clôture du contrat UREP08_VS est de 0,74\$.

Par ailleurs, par relation d'arbitrage, la somme des prix des contrats échangeables sur un marché donnée (par exemple, sur le *Vote Share Market*) doit être exactement égale à 1\$, si l'on exclut les coûts de transaction implicite. Cette relation tient grâce au fait que tout investisseur peut échanger à tout moment un portefeuille contenant une unité de chaque contrat échangeable sur le marché (par exemple, un contrat UDEM08_VS et un contrat UREP08_VS) contre 1\$ auprès de l'IEM. Or, si l'on regarde les prix de clôture publiés par l'IEM, cette relation ne tient

⁵ A notre connaissance, une telle situation ne s'est jamais produite dans l'histoire des élections présidentielles américaines, ce qui justifie que nous qualifions ce prix d'artificiellement élevé. Depuis la généralisation du scrutin populaire comme mode de désignation du collège électoral en 1836, le candidat Lyndon B. Johnson (Démocrate, élu Président en 1964) a obtenu le pourcentage de vote le plus important de tous les scrutins présidentiels, avec une part de 61,1% des suffrages exprimés.

pas toujours. En reprenant l'exemple précédent entre le 1^{er} et le 4 mai 2008, la somme des prix de clôture des contrats UDEM08_VS et UREP08_VS s'établit à 1,27\$ pour les trois premiers jours de la période et à 1,241\$ le 4 mai 2008. Cette règle de détermination des prix de clôture fait donc apparaître des anomalies temporaires, à de rares occasions. Néanmoins, le prix de clôture constitue une série continue de prix de marché, dont la pertinence est d'autant plus forte que la liquidité du marché est élevée. Par conséquent, dans les dernières semaines de la campagne présidentielle, traditionnellement la période où le marché est le plus liquide (mesuré par le volume quotidien échangé sur le marché), l'utilisation des prix de clôture ne pose pas de problème majeur. Pour éviter ces quelques anomalies très ponctuelles, nous aurions pu utiliser un prix synthétique calculé comme le milieu de la fourchette de prix, déterminé à partir des meilleures cotations *bid* et *ask* du carnet d'ordres, chaque soir à minuit. Néanmoins, cette information n'était pas immédiatement disponible auprès de l'IEM.

4.1.2. Des sondages nationaux hétérogènes

Nous avons choisi de constituer notre échantillon de sondages nationaux à partir de sondages d'opinion remplissant trois caractéristiques essentielles.

Tout d'abord, nous n'avons retenu que les sondages établis par des instituts renommés ou publiés par des médias de grande diffusion et à couverture nationale ou internationale. En effet, il existe un certain nombre de sondages, réalisés de façon plus ou moins formelle, et publiés par des organismes ou associations via leurs sites Internet ou des publications spécialisées. Pour ces sondages, il est très souvent difficile d'obtenir un compte-rendu détaillé des méthodes, questionnaires et des ajustements réalisés au cours de l'enquête. Nous estimons qu'il ne nous est pas possible de vérifier la fiabilité des informations publiées par ces organismes et la rigueur des méthodes employées, notamment dans le processus d'échantillonnage de la population (si toutefois il y a un réel processus de sélection des répondants). Nous faisons l'hypothèse que ces sondages, publiés le plus souvent de façon irrégulière et confidentielle, n'ont pas ou très peu d'influence sur les investisseurs actifs sur l'IEM.

Ensuite, nous avons sélectionné uniquement les organismes qui ont publié régulièrement des résultats de sondage au cours de la campagne électorale. Un certain nombre de médias, notamment des magazines et des journaux locaux, publient tout au plus un ou deux sondages durant toute la campagne électorale. Nous considérons que la publication de ces résultats est trop ponctuelle et trop aléatoire pour avoir un impact significatif sur le comportement des investisseurs. Toutefois, afin d'étudier un éventuel impact de cette catégorie de sondage, nous avons intégré à notre panel trois organismes de sondages qui n'ont publié qu'un nombre limité de résultats au cours de la campagne : GfK (3 sondages), Marist (4 sondages) et Time (3 sondages).

Les sondages sont, a priori, d'autant plus susceptibles d'avoir un impact sur la valorisation des contrats de l'IEEM qu'ils sont publiés par des entreprises relativement connues du public américain et international.

Enfin, dans un souci de fiabilité de l'information, tous les résultats de sondages présents dans notre panel sont issus directement des compte-rendu publiés par les instituts de sondage ou, dans les quelques cas où les compte-rendu n'étaient pas (ou plus) disponibles dans les archives des instituts, des données publiés par le New York Times dans son dossier de suivi quotidien des résultats des enquêtes d'opinion. Lorsque nous pouvions obtenir l'information de chacune des deux sources, nous avons pu vérifier que les données présentées par les deux sources concordaient exactement. L'unique différence réside dans le fait que les résultats présentés par le New York Times ne sont pas exhaustifs et ne garantissent pas que l'on prenne bien en compte l'intégralité des sondages publiés par un organisme. Toutefois, comme nous n'utilisons cette dernière source que pour vérifier les résultats d'un sondage dont nous avons eu connaissance par ailleurs, la question de l'exhaustivité ne s'est pas avérée problématique, à l'exception de quelques cas (moins d'une dizaine) pour lesquels nous avons eu recours à d'autres sources d'informations (le plus souvent, le site Internet *USA Election Polls* qui recense nombre de sondages publiés pendant la campagne électorale 2008). La liste des instituts de sondage retenus dans notre panel ainsi que le nombre de résultats d'enquête d'opinion utilisé pour chaque institut se trouve en annexe (Tableau 1).

Pour chacun des sondages, le panel recense à la fois des données relatives au résultat (pourcentages de voix exprimés pour le candidat démocrate, pour le candidat républicain, pour tout autre candidat et pourcentage d'indécis) et à la qualité de l'enquête elle-même (dates de début et de fin de l'enquête, date de publication des résultats, taille de l'échantillon, marge d'erreur associée aux résultats, institut en charge de l'enquête, sponsors de l'enquête). Enfin, lorsque l'information était disponible, la base de données contient également des données évaluant l'incertitude des intentions de vote exprimées par les répondants (sous la forme de proportion des répondants exprimant un soutien fort, léger ou une simple inclination pour leur candidat). Cette dernière donnée a pour but d'évaluer le niveau d'incertitude potentiellement retenu par les investisseurs au moment de prendre en compte le résultat d'une nouvelle enquête d'opinion.

4.1.3. Un large échantillon de sondages locaux

Du fait du caractère local des enquêtes d'opinion menées dans chaque Etat, les résultats de ces enquêtes ne sont pas toujours diffusés au niveau national, sauf lorsqu'ils concernent les principaux Etats du pays comme la Californie, le Texas ou l'Etat de New York. Pour réunir les résultats de sondages locaux, nous avons donc utilisé les informations fournies par le site

RealClearPolitics. Pour chacun des cinquante Etats, notre panel, dont une description est fournie en annexe (Tableau 2), regroupe des données similaires au panel national (dates de début et de fin d'enquête, institut en charge de l'enquête, taille de l'échantillon, marge d'erreur et pourcentages d'intentions de vote exprimés pour chacun des deux principaux candidats). Bien que notre recherche au niveau local porte sur l'impact des résultats des sondages dans les Etats, et tout particulièrement dans les *swing states*, nous n'avons pas souhaité limiter *a priori* la liste des Etats à intégrer dans notre panel. La définition des *swing states* est en effet sujette à controverse. S'il est certain que ces Etats sont identifiés comme ceux qui changent de couleur politique d'un scrutin à l'autre, il est difficile de prévoir *ex ante* les résultats des élections dans chaque Etat. D'autre part, la présence d'Etats pour lesquels l'issue du scrutin est relativement prévisible (Rhode Island ou l'Oklahoma par exemple) nous servira à constituer des variables test permettant de confirmer ou d'infirmier l'impact des *swing states*.

4.2. Analyse de l'échantillon

Dans les paragraphes qui suivent, nous menons une analyse quantitative et qualitative de notre échantillon. Nous présentons également la façon dont nous tentons d'homogénéiser l'information contenue dans notre base de données afin de pouvoir l'utiliser efficacement dans notre recherche.

4.2.1. Données chiffrées

Nous disposons des prix de marché (relevés à chaque clôture quotidienne) pour les marchés *Vote Share* et *Winner-Take-All* sur une période de 917 jours compris entre le 1^{er} juin 2006 et le 7 novembre 2008. Nous avons constitué un échantillon de 642 sondages nationaux publiés entre le 21 février 2007 et le 4 novembre 2008 par 21 organismes différents. Dans le cadre de notre objectif de comparaison des prévisions de marché avec les prévisions issues des sondages, les différences de caractéristiques temporelles entre les deux séries de données représentent un obstacle majeur. Les prix de marché constituent une série continue de données alors que les résultats des sondages sont publiés de façon ponctuelle et forment donc une série de données discontinue. Afin de limiter le décalage entre les deux séries de données, nous disposons de deux types de sondages : (1) les sondages classiques qui sont publiés à intervalles de temps plus ou moins réguliers et (2) les sondages de type « trackers » qui sont publiés quotidiennement et constituent de ce point de vue une série de données continue. Les *trackers* sont des sondages agrégés, le plus généralement créés à partir des résultats d'enquêtes d'opinion menées quotidiennement pendant les trois derniers jours précédant la publication des sondages. Sur les

642 sondages nationaux dont nous disposons, 463 sont des trackers publiés par Diageo (58 publications), Gallup (229 publications), Rasmussen (147 publications) et Zogby (29 publications). Grâce à Gallup, nous disposons ainsi d'une série temporelle continue de sondages publiés entre le 12 mars 2008 et le 2 novembre 2008⁶. Les 179 sondages qui complètent l'échantillon ont été publiés par des instituts et des sponsors de renom, parmi lesquels le *New York Times*, le *Washington Post*, ABC, CNN et Newsweek. La liste complète des instituts de sondage et des sponsors est disponible en annexe (Tableau 1).

Le panel de sondages locaux est constitué de 1 458 résultats d'enquêtes d'opinion réalisées le plus souvent par des organismes locaux. Certains d'entre eux sont officiellement affiliés au parti démocrate ou au parti républicain. Les principaux organismes ayant publié des sondages locaux durant la campagne électorale 2008 sont Rasmussen (361 publications), SurveyUSA (254), Research 2000 (77), Quinnipiac University (77) et CNN-Times (51). Aucun autre institut de sondage ne contribue à plus de 50 publications.

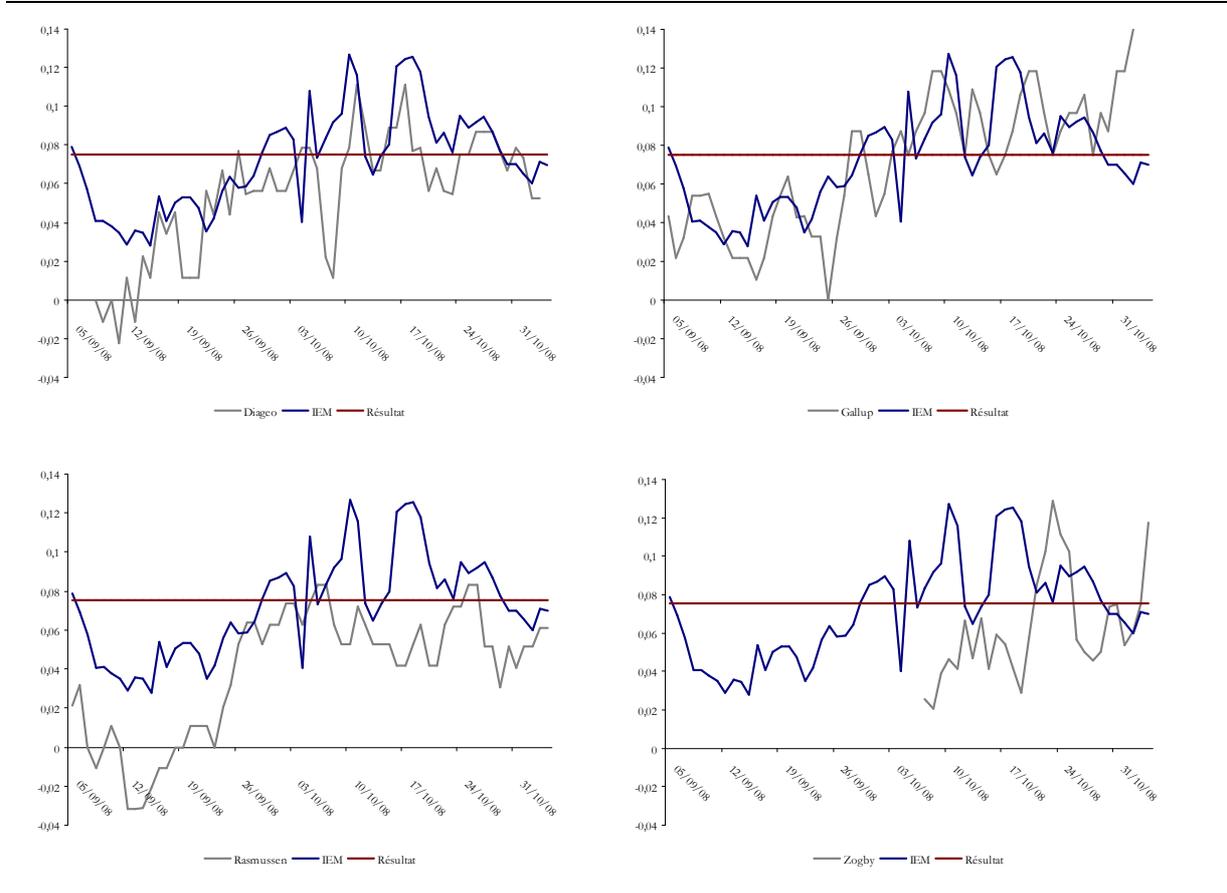
4.2.2. Volatilité et versatilité des sondages nationaux

Les sondages nationaux sont très hétérogènes. D'un organisme à l'autre, les formulations des questions ou le type de questions posées peuvent s'avérer très différents. Pour pouvoir agréger les données et les comparer avec les prix de marché, il nous était indispensable d'homogénéiser la présentation des résultats. Tout d'abord, nous répartissons les réponses des sondés en trois catégories distinctes en fonction de la préférence exprimée : intention de vote pour le candidat démocrate, intention de vote pour le candidat républicain et indécision. Les instituts de sondage diffèrent souvent sur le critère des indécis, certains offrant aux sondés la liberté d'exprimer une préférence pour un autre candidat, de nommer un candidat (cette option est souvent utilisée avant la nomination officielle des candidats par les deux principaux partis) ou de refuser d'exprimer leur opinion. Nous considérons que tous les répondants n'ayant pas apporté leur voix à l'un ou l'autre des deux principaux candidats sont des indécis. Dans le cadre de notre recherche, nous ferons abstraction de ces indécis pour deux principales raisons. Tout d'abord, une proportion significative des indécis n'ira pas voter le 4 novembre, leur indécision n'affectera donc pas le résultat de l'élection en tant que tel. D'autre part, lorsque les instituts de sondage cherchent à lever l'indécision en restreignant la liberté de réponse, nous remarquons que

⁶ Gallup n'a pas publié de tracker par 7 fois au cours de la période, en raison de jours fériés (Fête nationale du 4 juillet par exemple) ou d'événements extraordinaires (intempéries ayant interrompu les liaisons téléphoniques avec une partie du sud du pays fin juin 2008).

les indécis ont tendance à se répartir équitablement entre les deux candidats. Il est donc peu probable que l'opinion des indécis influence considérablement l'issue du scrutin.

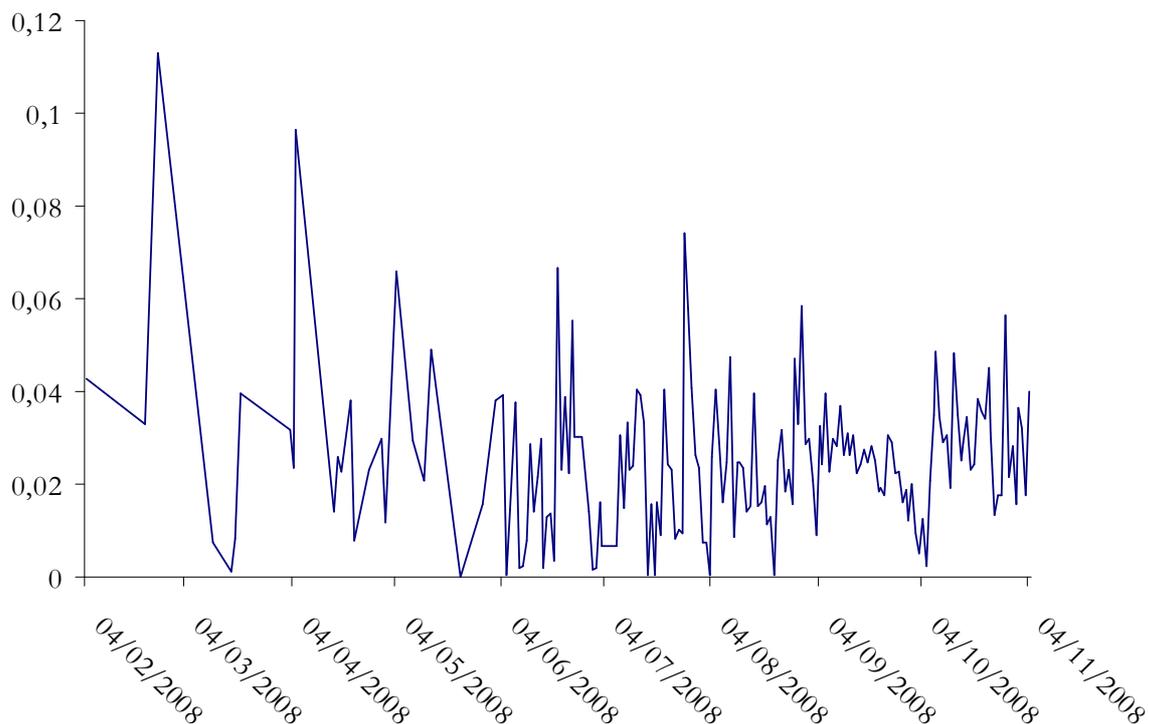
Graphique 1 : Comparaison des écarts normés des résultats des quatre Trackers, du prix de marché et du résultat des élections



L'analyse du panel des sondages montre une disparité conséquente des résultats. Sur le graphique 1⁷, nous observons que les résultats des enquêtes d'opinion sont en apparence beaucoup plus volatils que les cotations du marché. Les résultats des sondages sont en particulier beaucoup plus sensibles à l'effet d'annonce des conventions de chaque grand parti : ainsi le candidat démocrate a-t-il tendance à voir les intentions de vote augmenter en sa faveur au cours de la convention démocrate et dans les jours qui suivent cette réunion, alors que le candidat républicain observe le même phénomène durant l'organisation de la convention de son propre parti. De façon beaucoup plus surprenante, les sondages publiés à une date donnée peuvent rapporter des résultats extrêmement différents. Le graphique 2 présente l'écart-type quotidien calculé entre les différents sondages publiés au cours d'une journée (par rapport à un sondage de référence que nous calculons comme la moyenne des résultats des sondages publiés ce jour).

⁷ Le même type de représentation graphique, pour l'ensemble des principaux instituts de sondages de notre échantillon, est disponible en annexe (Graphique 6).

Graphique 2 : Ecart-type quotidien des résultats des sondages d'opinion



Certaines valeurs d'écart-type sont très élevées, soulignant la forte disparité des résultats publiés au cours d'une même journée. Plus la date de publication est éloignée de la date du scrutin, plus la probabilité de retrouver des valeurs élevées de l'écart-type est forte. A titre d'exemple, deux sondages publiés le 25 février 2008 prévoient la victoire d'un candidat différent, l'un des résultats annonçant un écart de 12 points entre les deux candidats. Le 5 avril 2008, Gallup prévoit une égalité parfaite entre les deux candidats alors que Diageo accorde une avance de 12 points au candidat démocrate. A mesure que l'on approche de la date du scrutin, la valeur de l'écart-type diminue et semble majoritairement influencée par l'importance accordée par les différents organismes de sondage aux indécis. Par exemple, l'écart-type des résultats des sondages du 28 octobre 2008 est plus élevé que la moyenne des écart-types de cette période, notamment du fait que la proportion d'indécis varie de 3% à 12% suivant les instituts de sondage. Néanmoins, les six sondages publiés à cette date prévoient tous la victoire du candidat démocrate, avec une avance relativement constante d'un organisme de sondages à l'autre.

4.2.3. Une inégale répartition des sondages locaux

Les Etats fédérés ne sont pas tous égaux quant à l'attention médiatique qui leur est dévolue au cours de la campagne électorale des élections présidentielles. Cette différence de traitement résulte principalement du mode de scrutin – suffrage universel indirect – utilisé lors de

ces élections. Contrairement à ce qui se passe en France par exemple, le Président n'est pas élu directement par les citoyens, mais par un collège électoral formé des représentants élus par les citoyens de chaque Etat. A l'exception du Maine (depuis 1972) et du Nebraska (depuis 1992), tous les Etats utilisent le système du *Winner-Take-All* selon lequel le candidat qui obtient la majorité des suffrages exprimés dans un Etat remporte l'ensemble des voix des représentants de cet Etat. De facto, les Etats possédant le plus de représentants au sein du collège électoral (les Etats les plus peuplés, le nombre de représentants étant fonction du poids démographique de l'Etat au sein de la fédération, tout Etat ayant au minimum trois représentants) sont ceux qui focalisent le plus les médias et, par voie de conséquence, ceux qui intéressent le plus les instituts de sondages. C'est ce qui explique, en particulier, que la Floride, la Pennsylvanie et l'Ohio, respectivement au quatrième, cinquième et sixième rang par leur nombre de représentants dans le collège électoral possède chacun environ 90 résultats de sondage dans notre échantillon. Parallèlement, des Etats tels que le Vermont (3 représentants), Rhode Island, l'Idaho ou Hawaï (4 représentants chacun) possèdent tous moins de cinq sondages dans l'échantillon.

Toutefois, le nombre de sondages locaux s'explique bien plus encore par l'incertitude relative portée au scrutin dans chaque Etat. Il existe bien des exemples où cette incertitude est très limitée : depuis la victoire du démocrate Lyndon B. Johnson en 1964, aucun candidat démocrate ne s'est plus jamais imposé dans le Wyoming, l'Oklahoma, l'Utah et l'Idaho. Pour nombre d'analystes, l'issue du scrutin dans ces Etats ne fait aucun doute et la contribution de ces Etats au panel de sondage est très faible (respectivement, 7, 9, 10 et 4 sondages). Le même phénomène se produit pour des Etats réputés démocrates : New York (seulement 31 sondages alors que cet Etat se classe troisième par son nombre de représentants), le Massachusetts (21 sondages), Rhode Island et Hawaï n'ont plus voté républicain depuis la réélection de Ronald Reagan en 1984. A l'inverse, les Etats où le scrutin s'annonce le plus serré, et notamment les Etats qui risquent de changer de couleur politique par rapport au précédent scrutin (les *swing states*) sont beaucoup plus médiatisés. Après avoir analysé notre échantillon de données, nous présentons les principaux résultats de notre recherche.

5. Résultats de la recherche

Notre recherche s'articule autour de trois questions fondamentales. Tout d'abord, nous essayons d'évaluer la performance des marchés de l'PIEM, à la fois en terme absolu (c'est-à-dire en comparant les prévisions du marché directement au résultat de l'élection) et en terme relatif (en comparant le prix de marché aux résultats des sondages de notre base de données). Nous

évaluons cette performance à divers horizons de temps, et en utilisant parfois plusieurs méthodes (notamment dans le cas de la performance relative). Nous proposons ensuite de vérifier si les résultats des sondages nationaux ont un impact sur les prix de marché. Enfin, nous tentons de répliquer les prix de marché à partir des résultats des sondages locaux, parus dans chaque Etat.

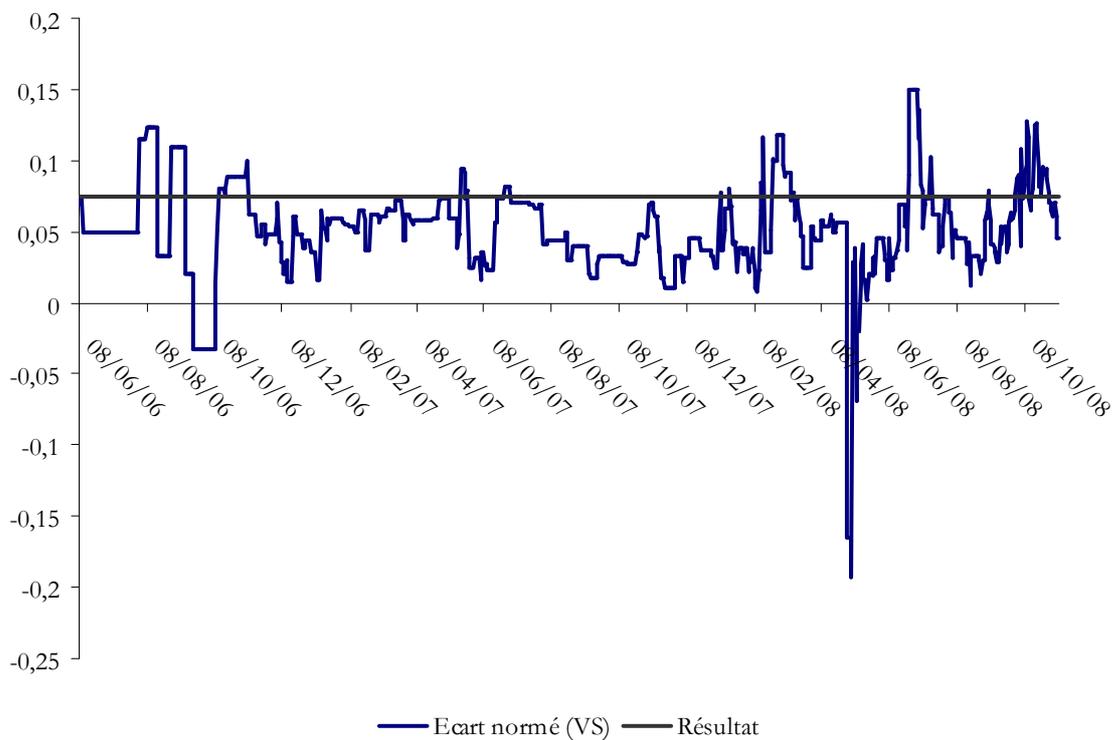
5.1. En 2008, les marchés ont-ils prévu le résultat du scrutin avec précision ?

Nous analysons tout d'abord la qualité de la prévision du marché en terme absolu et de proposer des explications à cette performance du marché.

5.1.1. Qualité de la prévision à court, moyen et long termes

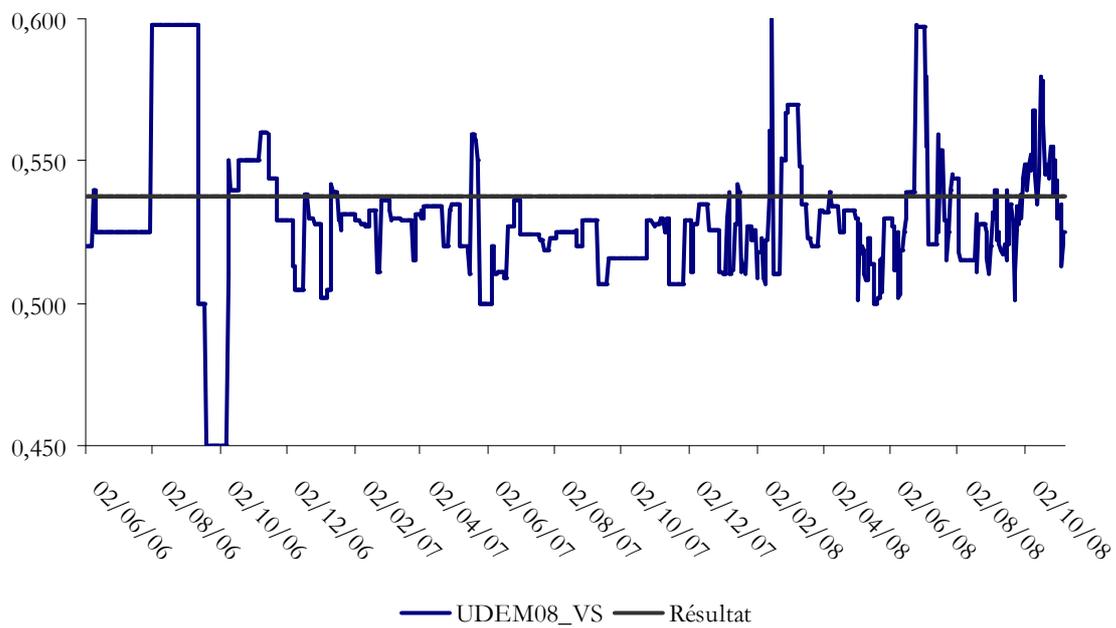
Lorsqu'on évalue la précision de la prévision du marché, il est courant de se focaliser sur la qualité de la prévision à très court terme, en comparant le prix de marché relevé la veille du scrutin au résultat de l'élection. Au soir du 3 novembre 2008, le *Vote Share Market* prévoyait un pourcentage des suffrages exprimés de 53,5% pour le candidat démocrate et 46,4% pour le candidat républicain. Après décompte des voix, la commission électorale a attribué 53,76% des voix au candidat démocrate et 46,24% des voix au candidat républicain (lorsque seules les voix exprimées pour l'un des deux principaux candidats sont prises en compte). Nous observons donc une erreur de 0,26% pour le contrat démocrate et 0,16% pour le contrat républicain. En arrondissant ces chiffres, il convient de rappeler qu'ils représentent, sur le marché, un écart de respectivement 0,3 cents et 0,2 cents quand le plus petit incrément de prix autorisé sur le marché est fixé à 0,1 cent. La performance du marché est donc particulièrement remarquable. Ces résultats ne sont d'ailleurs pas uniques et sont observés de manière plus ou moins régulière à chaque élection présidentielle américaine, comme le relèvent Berg, Forsythe et Rietz (1997, pp. 220-221).

Graphique 3 : Evolution de l'écart normé sur le *Vote Share Market* par rapport au résultat de l'élection

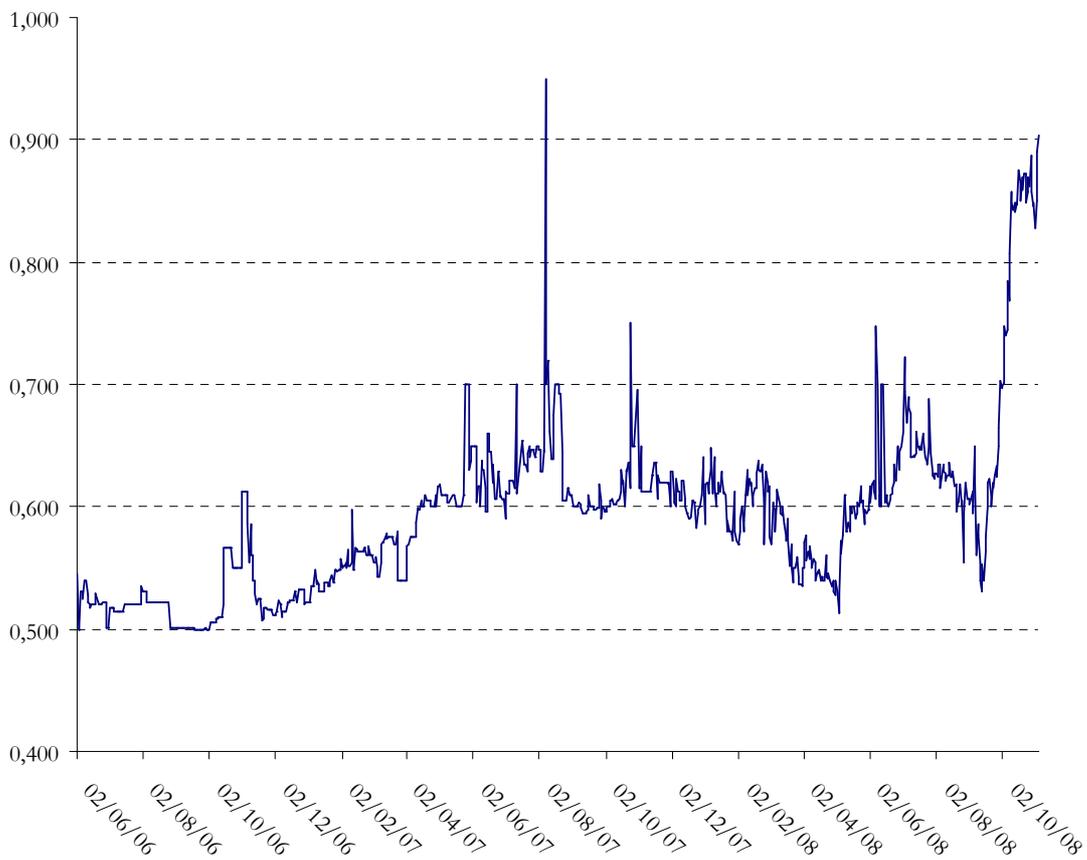


Tentons maintenant d'élargir l'horizon de prévision progressivement pour analyser la stabilité de la performance du marché à plus long terme. Le graphique 3 montre un comparatif entre les cotations du marché (mesurées ici par l'écart entre le prix du contrat démocrate et celui du contrat républicain, normé par la somme des prix des deux contrats) et le résultat de l'élection (mesuré de la même façon, avec les pourcentages des suffrages exprimés pour chaque candidat). Au cours de la semaine précédant le scrutin, la moyenne des cotations du contrat démocrate s'établit à 0,538\$, soit presque exactement le résultat de l'élection, et celui du contrat républicain à 0,467\$. La performance reste remarquable lorsqu'on élargit l'horizon à un mois (0,550\$ en moyenne pour le contrat démocrate et 0,461\$ pour le contrat républicain) ou à deux mois (0,540\$ pour le contrat démocrate et 0,467\$). Sur ces deux périodes, la moyenne des erreurs absolues du marché par rapport au résultat du scrutin s'établit respectivement à 1,4% et 1,3% pour le contrat démocrate et, 0,9% et 1,1% pour le contrat républicain. Sur l'ensemble de la période d'ouverture, la moyenne des erreurs absolues s'élève à 1,9% pour le contrat UDEM08_VS et à 1,87% pour le contrat UREP08_VS.

Graphique 4 : Evolution du prix du contrat UDEM08_VS par rapport au résultat de l'élection



Graphique 5 : Evolution du prix du contrat DEM08_WTA



Si l'on considère maintenant les prix des contrats échangés sur le *Winner-Take-All Market* (graphique 5), on remarque avec surprise que le marché n'a jamais cru en une victoire du candidat républicain. Alors que nombre de sondages ont annoncé périodiquement une victoire possible du candidat républicain, notamment pendant la convention républicaine de septembre 2008, le marché a toujours accordé au candidat démocrate une probabilité de victoire supérieure à 50%. Nous comprenons également à la vue de ce graphique pourquoi le *Winner-Take-All Market* est celui qui attire le plus les investisseurs. C'est en effet le marché sur lequel les prix sont les plus volatils et où les investisseurs pouvaient acquérir des contrats DEM08_WTA à 0,55\$, par exemple, et recevoir 1\$ s'ils les conservaient jusqu'à l'échéance. Il est presque impossible de réaliser un profit de 0,45\$ sur un seul contrat sur le *Vote Share Market*. Nous présentons par la suite nos principales conclusions sur la performance absolue du marché.

5.1.2. Eléments d'explications et conclusions

Même sur très longue période, parfois plus de deux ans avant le jour de l'élection, le marché réussit donc à prévoir de façon relativement précise l'issue du scrutin. La précision de la prévision augmente lorsqu'on réduit l'horizon de temps considéré. On aurait pu s'attendre à une convergence des prix du marché relativement progressive vers le résultat de l'élection. En réalité, nous pouvons remarquer sur le graphique que le marché a eu tendance à sous-estimer le pourcentage des votes attribué au candidat démocrate pendant une période de temps relativement longue, puis inversement à surestimer le score du candidat démocrate pendant les dernières semaines de la campagne électorale.

Pourtant, nous ne pouvons nous empêcher de rester perplexe devant cette performance du marché. Si des facteurs tels que l'instabilité boursière, le climat économique, le choix du futur vice-président par chacun des deux candidats a une réelle influence sur l'issue de l'élection, comme le suggèrent nombre de journalistes, comment est-il possible que le marché prenne en compte ces facteurs si longtemps à l'avance ? Quel investisseur pouvait prévoir, au milieu de l'année 2006, la panique boursière du mois de septembre 2008 ? Qui pouvait assurer que Barack Obama serait le candidat du parti démocrate quand la nomination d'Hillary Clinton était encore quasi certaine ? Qui avait envisagé la nomination de la gouverneure de l'Alaska, Sarah Palin, au sein du ticket républicain ? Pourtant, si l'on croit les résultats des sondages, toutes ces informations étaient de nature à influencer le vote des citoyens américains. Gelman et King (1993) formulent une théorie en réponse à cette argument : le déroulement de la campagne électorale n'a aucune influence sur le vote. Cette théorie suggère que le déroulement de la campagne électorale a peu d'importance et que le résultat du scrutin serait plus ou moins prévisible grâce à certains indicateurs politiques, économiques et sociologiques. Elle

s'accompagne généralement du développement de modèles prédictifs basés sur des variables telles que le taux de croissance de l'économie entre deux élections ou la lassitude de l'opinion vis-à-vis du parti au pouvoir. Ces modèles semblent avoir donné des résultats relativement fiables au cours des dernières décennies, comme le montrent en particulier Leigh et Wolfers (2006) dans le cas des élections générales australiennes de 2004. Après avoir étudié la performance absolue du prix de marché, nous proposons d'étudier cette performance relativement à celle établie par un autre élément de prévision, les sondages d'opinion.

5.2. Le pouvoir de prédiction du marché est-il supérieur à celui des sondages d'opinion ?

5.2.1. Méthodologie

Nous utilisons dans cette section les résultats des enquêtes d'opinion dans le but de les comparer aux cotations relevées sur le *Vote Share Market*. Pour chaque sondage, nous relevons les préférences exprimées par *likely voters* (c'est-à-dire le groupe de répondants ayant indiqué qu'ils avaient l'intention de se rendre aux urnes le jour du scrutin) ou, à défaut, par les *registered voters* (c'est-à-dire le groupe de répondants ayant indiqué qu'ils étaient inscrits sur les listes électorales, indépendamment du fait qu'ils aillent voter ou non le jour du scrutin). Sur les 642 sondages nationaux que compte notre panel, 547 prennent en compte l'opinion des *registered voters*. La première catégorie de répondants est davantage représentative de la population qui votera effectivement le 4 novembre, les intentions de vote exprimées par cette catégorie doivent donc logiquement prédire l'issue du scrutin avec une meilleure précision. Afin de tester la meilleure performance du marché par rapport aux enquêtes d'opinion, nous utilisons une formule d'agrégation des prix (ou des pourcentages d'intentions de vote) des deux principaux candidats. Cette formule, qui renvoie au concept de *marginalized lead*, est notamment utilisée par Berg, Nelson et Rietz (2008). A partir des prix de marché relevés quotidiennement sur les contrats UDEM08_VS et UREP08_VS, nous calculons l'écart des prix entre le contrat du candidat démocrate (P_{DEM}^{VS}) et le contrat du candidat républicain (P_{REP}^{VS}), normé par la somme des deux prix des contrats selon la formule suivante :

$$E^{VS} = \frac{P_{DEM}^{VS} - P_{REP}^{VS}}{P_{DEM}^{VS} + P_{REP}^{VS}} \quad (3)$$

La normalisation par la somme des prix de marché des deux contrats permet d'allouer proportionnellement aux prix des contrats les écarts de prix par rapport à la règle d'arbitrage

selon laquelle la somme des prix des contrats doit toujours être égale à \$1 (nous avons développé un exemple de telle déviation de prix à la section 4.1.1).

De la même façon, à partir des pourcentages d'intentions de vote exprimées par les répondants aux sondages, nous calculons l'écart entre le pourcentage de voix reçu par le candidat démocrate (P_{DEM}) et le pourcentage de voix reçu par le candidat républicain (P_{REP}), normé par la somme des pourcentages reçus par chacun des deux candidats, selon la formule suivante :

$$E^{Sondage} = \frac{P_{DEM} - P_{REP}}{P_{DEM} + P_{REP}} \quad (4)$$

La normalisation par la somme des intentions de vote reçues par chacun des deux principaux candidats permet d'allouer les votes en faveur d'autres candidats ou les votes des indécis proportionnellement aux intentions de vote déjà exprimées (nous avons notamment développé les implications de ce mécanisme à la section 4.2.2).

Enfin, nous procédons de même avec les résultats du scrutin du 4 novembre 2008, en calculant l'écart entre le nombre de voix reçu par le candidat démocrate (v_{DEM}) et le nombre de voix reçu par le candidat républicain (v_{REP}) le jour de l'élection. Les suffrages exprimés en faveur d'autres candidats et les votes blancs ou nuls sont alloués proportionnellement aux votes déjà exprimés en faveur de chacun des deux candidats, en normalisant l'écart précédemment obtenu par la somme des voix obtenues par chacun des deux candidats :

$$E^{Résultat} = \frac{v_{DEM} - v_{REP}}{v_{DEM} + v_{REP}} \quad (5)$$

A partir de ces écarts, nous calculons, pour chaque jour représenté dans notre base de données, la distance (erreur absolue de prévision) entre, d'une part, les prévisions du marché et le résultat de l'élection et, d'autre part, les prévisions issues des sondages et le résultat de l'élection :

$$D^{VS} = |E^{VS} - E^{Résultat}| \quad (6.1)$$

$$D^{Sondage} = |E^{Sondage} - E^{Résultat}| \quad (6.2)$$

Connaissant ces distances, nous estimons que le marché produit une meilleure prévision du résultat à l'échéance lorsque la valeur de D^{VS} est inférieure à la valeur de $D^{Sondage}$. Afin d'enregistrer la performance de chacune des deux séries de données, nous créons une variable binomiale qui prend la valeur 1 lorsque le marché produit une meilleure prévision, et 0 autrement. Enfin, nous exécutons un test binomial sur cette variable.

Nous réalisons ce test sur trois séries de données issues de notre base de sondages. Tout d'abord, nous comparons simplement les prix de marché le jour de la publication d'un sondage au résultat du sondage. Pour cette comparaison, nous sommes contraints de faire l'hypothèse que

les investisseurs n'ont pas accès à l'information issue des sondages. Ensuite, considérant que bien que les résultats soient publiés à une certaine date ils ne représentent que l'opinion de l'échantillon interrogé au cours de l'enquête, nous comparons le résultat du sondage au prix de marché qui était en vigueur le dernier jour de l'enquête terrain réalisée par l'institut de sondages. Enfin, nous pensons également que plusieurs sondages publiés le même jour ne doivent pas s'apprécier individuellement mais plutôt par la précision que le groupe de sondages publié au cours d'une même journée est capable d'apporter. Aussi, pour chaque jour au cours duquel au moins un sondage est publié, nous faisons la moyenne des distances calculées pour chaque sondage et nous comparons cette moyenne au prix de marché en vigueur le jour de la publication des résultats des sondages. Nous présentons dans la partie suivante les principaux résultats de nos tests.

5.2.2. Quand le marché surpasse les sondages

Lorsque l'on considère l'ensemble des sondages disponibles sur la période d'ouverture du marché, quelle que soit la méthode de comparaison utilisée, la prévision établie à partir des prix de marché est plus précise que celle dérivée des résultats des sondages avec une fréquence oscillant entre 66% et 70%. Ces résultats sont semblables à ceux établis lors d'études sur les élections présidentielles américaines précédant celle de 2008, et notamment dans l'étude de Berg, Nelson et Rietz (2008). Néanmoins, lorsque l'on détaille les résultats par période, il apparaît nettement que le marché surpasse beaucoup plus souvent les sondages lorsque l'on se situe à plus d'un mois de l'échéance électorale. Au cours du mois de septembre 2008, le marché obtient ainsi un taux de réussite oscillant entre 79% et 87%, suivant les méthodes de comparaison utilisée. Le tableau 1 montre d'ailleurs que, dans les périodes temporelles à plus d'un mois de l'élection, la précision du marché est relativement comparable d'une méthode de comparaison à l'autre. En revanche, les scores obtenus par le marché sont moins significatifs au cours du mois d'octobre et dans les cinq derniers jours précédant l'élection.

Le fait de comparer le résultat d'un sondage au prix de marché relevé le dernier jour de l'enquête améliore la performance des sondages par rapport au cas où nous comparons le résultat du sondage au prix de marché en vigueur le jour de sa publication. Le taux de réussite des sondages au test binomial augmente en moyenne d'environ 3%, avec une hausse beaucoup plus marquée au cours des mois de septembre et d'octobre. Dans les cinq derniers jours précédant le scrutin, la surperformance du marché n'est statistiquement significative qu'à un niveau de 11%.

L'utilisation d'un sondage moyen par jour de publication donne des résultats sensiblement différents. Si la performance du marché reste constante sur l'ensemble du panel et à plus de 100 jours de l'élection, elle s'améliore nettement à plus d'un mois et plus de deux mois de l'élection

(+8 points dans chacun des deux cas, par rapport à la seconde méthode de comparaison). Cela signifie que sur cette période, si quelques sondages parviennent ponctuellement à produire des prévisions plus précises que celles du marché, en moyenne, les prévisions établies par les sondages sont très éloignées (ou tout au moins relativement plus éloignées que celles du marché) du résultat final de l'élection. En revanche, le sondage moyen se révèle beaucoup plus précis que les prix de marché lors du mois précédant la date du scrutin. Entre le 4 octobre 2008 et le 4 novembre 2008, le sondage moyen représente une meilleure prévision du résultat par 21 fois sur 32, soit un score de 66%. Même si cette surperformance n'est pas statistiquement significative, notamment en raison de la faible taille de l'échantillon, il n'en reste pas moins que ces résultats vont à l'encontre des autres résultats que nous avons obtenus jusqu'à présent. Au cours des cinq derniers jours précédant l'élection, le marché n'est plus précis qu'une seule fois. Toutefois, il faut noter que cette meilleure précision du marché a lieu précisément la veille de l'élection. Aussi, lorsque l'on compare la performance du marché à la performance des sondages à très court terme, on relève très justement que le marché est plus précis que les instituts de sondage, ce que ne manque pas de souligner le service de presse de l'Université de l'Iowa (UI News Services, 24 novembre 2008). Nous proposons dans la partie suivante des explications potentielles à cette contre-performance du marché.

5.2.3. L'année 2008 n'est pas propice au marché

Les résultats sont en apparence discordants. Si nous observons que le fait de comparer le résultat du sondage au prix en vigueur sur le marché le dernier jour de l'enquête et non le jour de la publication des résultats de l'enquête n'augmente pas significativement la performance des sondages, la comparaison des prix de marché aux sondages moyens publiés quotidiennement porte à l'avantage des sondages, au moins à court terme. Nous pensons que cette dernière méthode est celle qui devrait être retenue. Berg, Nelson et Rietz (2008) ont également proposé d'utiliser une moyenne des résultats des sondages comme élément de comparaison aux prix de marché, en retenant les résultats des cinq derniers sondages publiés dans le calcul de la moyenne. Si l'objectif est de chercher à comparer le marché et les sondages dans le but d'établir l'existence (ou non) d'une relation de causalité entre les deux types de données, il nous semble que la moyenne des cinq derniers sondages n'est pas adaptée à notre cas. Ce système impose une certaine inertie au sondage lorsqu'il y a peu de publications de résultats sur une période. Par ailleurs, si nous n'utilisons qu'un seul prix de marché par jour de cotation, nous estimons plus juste de n'utiliser qu'un résultat agrégé de sondages par jour de publication. Dans le cas contraire, si nous comparons systématiquement chaque sondage au prix du marché, pourquoi ne pas comparer chaque sondage à chacun des prix établis par le marché au fur et à mesure de la

journée, comme s'il s'agissait d'une publication de prévision par un groupe d'investisseurs ? Toutefois, le sondage agrégé n'est pas non plus la solution idéale car le prix de marché n'est pas un prix moyen mais un prix d'équilibre établi entre offreurs et demandeurs de liquidité.

Plusieurs facteurs peuvent expliquer la mauvaise performance relative du marché sur les derniers jours de la campagne électorale. Dans notre panel, la majorité des résultats de sondage basés sur les *likely voters* est concentrée dans le mois précédant l'élection. Or, Crespi (1988) affirme que ces sondages sont plus précis que ceux qui sont basés sur les *registered voters* (et qui forme l'essentiel de notre panel). Le tableau 3 de l'annexe montre que sur les trois périodes où le marché prévoit l'issue du scrutin de façon plus précise que les sondages, le panel est constitué à plus de 90% de sondages basés sur les *registered voters*. En revanche, sur les deux périodes précédant le scrutin, cette proportion est réduite à 62,6% au mois d'octobre et 56% au mois de novembre. Pour 28 des 642 sondages de notre panel, nous disposons à la fois des résultats basés sur les *registered* et les *likely voters*. Pour ceux-ci, nos résultats sont basés sur les intentions de vote exprimées par les *likely voters*. Si toutefois nous avions utilisé les intentions de vote exprimées par les *registered voters*, la proportion de sondages utilisant les résultats des votes des *registered voters* passerait à 79% pour le mois d'octobre et 76% pour le mois de novembre. Sur les cinq derniers jours de la campagne, les prix de marché se seraient révélés plus précis dans trois cas sur cinq. Entre 6 et 31 jours avant l'élection, le score des prix de marché passerait à 41%. Les résultats de nos tests pourraient donc s'expliquer par des résultats plus précis parce qu'ils prennent plus en compte les intentions de vote exprimées par les *likely voters*. Toutefois, comme nous ne disposons pas d'informations concernant ce caractère pour les études précédentes, nous ne pouvons pas vérifier cette hypothèse. Par ailleurs, la précision des prévisions issues des enquêtes d'opinion est généralement d'autant plus importante que l'incertitude sur l'issue du scrutin est faible. De façon conjoncturelle, le fait que l'élection du candidat démocrate était très probable même quelques semaines avant le scrutin a certainement joué en faveur de la précision des sondages.

Forsythe et Rietz (1997) montrent que la prévision des marchés de l'IEM est d'autant plus précise que le volume échangé (mesuré en \$) au cours de la campagne électorale est élevé. Le tableau 4, disponible en annexe, montre que le volume échangé sur le *Vote Share Market* entre le 1^{er} juin 2006 et le 7 novembre 2008 est extrêmement faible (10 012\$). C'est le second plus faible volume échangé sur le *Vote Share Market* pour une élection présidentielle après l'édition de 1996⁸ (3 628\$ d'après Berg, Nelson et Retz (2003)). Il est possible d'imaginer que la faible incertitude sur l'issue du scrutin puisse justifier le faible intérêt des investisseurs pour le marché. Le marché

⁸ Nous ne prenons pas en compte ici l'élection de 1988 où seul un nombre limité de participants avait accès à l'IEM. Cette année-là, 8 123\$ ont été échangés d'après Berg, Nelson et Retz (2003).

prévoit en effet une victoire du candidat démocrate depuis le mois de mai 2008. Depuis cette date, les cotations du marché n'ont jamais été en faveur du candidat républicain. D'autre part, il ne faut pas oublier que les deux derniers mois de la campagne électorale sont marqués par une période de forte volatilité des marchés financiers traditionnels. Même si l'on suppose que l'IEM est décorrélé des marchés financiers et que les montants engagés sur l'IEM par un investisseur ne font pas partie de son portefeuille d'investissements, nous ne pouvons négliger le fait que l'attention des investisseurs ait pu être détournée de l'IEM pour se focaliser sur les marchés traditionnels.

5.2.4. Conclusion

Contrairement aux résultats des précédentes études, notre échantillon montre une meilleure précision des prévisions issues des sondages agrégés en moyenne quotidienne dans les dernières semaines de la campagne électorale. Ce résultat peut être à la fois imputable à une meilleure performance des sondages, dans un contexte de faible incertitude de l'issue du scrutin, et à une moindre performance du marché. Certains investisseurs ont en effet pu délaisser les marchés de l'IEM pour se concentrer sur les problèmes provoqués par l'instabilité des marchés financiers traditionnels. De même nous ne pouvons pas exclure que le marché ait surréagi aux nouvelles publiées lors des deux derniers mois précédant le scrutin. En particulier, on observe que le marché a fortement décoté le contrat démocrate après la convention républicaine et la nomination de Sarah Palin au sein du ticket républicain. De même, le prix du contrat démocrate apparaît largement surévalué à partir de la fin du mois de septembre, alors même que le candidat républicain vient d'interrompre sa campagne pour travailler sur la crise financière et envisage de ne pas participer au premier débat présidentiel. Quoi qu'il en soit, la relation entre le prix de marché et les résultats des enquêtes d'opinion apparaît complexe et il nous semble intéressant d'étudier la question de l'influence des résultats des sondages dans la construction des prix des contrats sur le *Vote Share Market*.

5.3. Les sondages influencent-ils les prix de marché ?

Nous tâchons maintenant de déterminer si les résultats des sondages nationaux ont influencé sur les prix du *Vote Share Market*. Pour cela, nous utilisons dans un premier temps le modèle proposé par Forsythe, Nelson, Neumann et Wright (1992). Nous proposons ensuite de construire un autre modèle par régression linéaire en considérant de nouvelles variables explicatives.

5.3.1. Le modèle de Forsythe, Nelson, Neumann et Wright (1992)

5.3.1.1. Présentation du modèle

La majorité des études menées sur le lien entre les prix de marché et les résultats des sondages a conclu à une absence de corrélation entre les deux séries de données. Dans leurs travaux publiés en 1992, Forsythe, Nelson, Neumann et Wright ont testé un modèle explicatif non linéaire des prix du *Vote Share Market* basé sur des données historiques et des hypothèses de relations de causalités formulées *a priori*. Nous proposons, dans un premier temps, de tester le même modèle sur notre base de données de l'élection présidentielle de 2008. Les auteurs supposent que le prix de marché devrait être égal à l'espérance du prix de marché en l'absence de nouvelle information, à laquelle ils ajoutent un facteur qui prend en compte la « surprise » créée par la publication d'un sondage, en mesurant cette « surprise » par la différence entre le résultat du sondage publié et les anticipations du marché concernant ce résultat. Les anticipations du marché concernant le résultat du sondage sont supposées être fonction des résultats du dernier sondage publié par le même institut de sondage et d'une mesure de la prise en compte de nouvelles informations par le marché depuis cette dernière publication (différence entre le dernier prix de marché et le prix de marché relevé lors de la dernière publication des résultats du sondage). Une variable muette vient annuler le terme correspondant à un institut de sondage si cet institut ne publie pas de nouveaux résultats dans le cours de la journée. Ce modèle est résumé par les trois formules ci-dessous :

$$M_t = E_t(M_t / \phi_{t-1}) + \sum_i \beta_i [P_{i,t} - E_t(P_{i,t} / \phi_{t-1})] D_{i,t} + \varepsilon_t \quad (7)$$

$$E_t(M_t / \phi_{t-1}) = \alpha_0 + \alpha_1 M_{t-1} + \alpha_2 M_{t-2} \quad (8)$$

$$E_t(P_{i,t} / \phi_{t-1}) = \gamma_{i,0} + \gamma_{i,1} P_{i,t-s} + \gamma_{i,2} (M_{t-1} - M_{t-s}) \quad (9)$$

Dans ces équations, M_t représente l'écart de prix sur le *Vote Share Market* à une date t tel que défini par la formule (3), ϕ_{t-1} représente toute l'information connue à la date $t-1$ et $P_{i,t}$ renvoie au résultat d'un sondage publié par l'institut i à la date t . Ce résultat est agrégé selon la méthode développée à la section 5.2.1. et définie par la formule (4). Les coefficients α , β et γ sont obtenus par une régression non linéaire. D'après ce modèle, nous pouvons conclure que les résultats des sondages influencent les prix de marché si les coefficients β sont significativement différents de zéro.

5.3.1.2. *Le modèle est invalidé*

Nous commençons par tester le modèle sur les quatre instituts de sondage qui ont fourni des *trackers* quotidiens dans les derniers jours ou les dernières semaines de la campagne électorale : Diageo, Gallup, Rasmussen et Zogby. Nous pensons en effet que ces instituts sont susceptibles d'avoir le plus fort impact sur les prix de marché, étant donné leur fréquence de publication de nouveaux résultats. Le tableau 5 de l'annexe montre les résultats de la régression non linéaire basée sur le modèle de Forsythe, Nelson, Neumann et Wright (1992). Lorsque nous testons le modèle sur l'ensemble de la période d'ouverture du marché (du 1^{er} juin 2006 au 7 novembre 2008), seuls les coefficients α_1 et α_2 apparaissent significativement différents de zéro. A la vue des résultats, nous ne pouvons rejeter l'hypothèse nulle selon laquelle les coefficients β de nos quatre *trackers* seraient nuls. Ces premiers résultats semblent donc mettre en évidence une absence de relation entre les prix de marché et les résultats des sondages. Cependant, il faut remarquer que nous ne disposons d'aucune publication de résultat de *tracker* avant le 12 mars 2008 (date de publication du premier *tracker* par Gallup). Avant cette date, nous forçons le modèle à comparer des prix de marché (non nuls) avec des variables toutes égales à zéro (mise à part les variables de prix de marché intégrées au modèle). Il est donc possible que la performance du modèle soit altérée au cours de la période où aucun résultat de sondage n'est publié. Nous réitérons l'expérience par quatre fois en définissant le début de la période de test par la première date de publication de chaque institut, soit le 12 mars 2008 pour Gallup, le 8 juin 2008 pour Rasmussen, le 11 septembre 2008 pour Diageo et le 7 octobre 2008 pour Zogby. Nous relevons essentiellement les mêmes résultats sur cette nouvelle série de tests. Sur ces périodes réduites, le coefficient α_2 n'est plus statistiquement différent de zéro au niveau de confiance de 5%. Entre le 8 juin 2008 et le 7 novembre 2008, l'intervalle de confiance à 95% du coefficient β de Rasmussen ne contient pas la valeur zéro. Cependant, la valeur estimée de ce coefficient est elle-même très proche de zéro et ce résultat semble bien plus le fruit du hasard que d'une relation de cause à effet. En effet, si l'on réduit la période d'étude, on observe la disparition de cette significativité dans les deux sous-périodes suivantes. Enfin, malgré l'absence de significativité de l'ensemble des variables de sondage, le modèle extériorise un niveau de R^2 relativement élevé. Nous envisageons deux explications à ce phénomène. Tout d'abord, sur l'ensemble de la période d'ouverture du marché, les premiers mois d'ouverture s'accompagnent d'une activité faible du marché (mesuré en terme de volume échangé) : les variations quotidiennes des prix de marché sur cette période sont donc relativement limitées, ce qui explique que l'on puisse prédire une grande partie des variations du prix de marché en utilisant le prix des deux jours précédents. De fait, lorsque l'activité du marché augmente, le prix de clôture relevé deux jours auparavant n'est plus

utilisé par le modèle. Néanmoins, il serait vain de croire que l'on peut modéliser le prix de marché uniquement en fonction du prix relevé un jour plus tôt à la clôture. A titre d'exemple, nous avons fait l'exercice sur la dernière période (du 7 octobre au 7 novembre 2008). Le modèle déterminé par la régression est le suivant : $M_t = 0,016 + 0,802 \cdot M_{t-1} + \varepsilon_t$. Sur un mois, l'erreur absolue moyenne de ce modèle est de 12,8%. Inutile de préciser qu'investir sur la base de ce modèle aurait pu s'avérer rapidement préjudiciable pour l'investisseur.

Toutefois, il est possible que nous ayons omis une partie importante de l'information en ne considérant que quatre instituts de sondage. Pour vérifier l'impact de cette hypothèse, nous ajoutons l'ensemble des instituts de sondage qui constituent notre base. Etant donné les limites imposées par la complexité du modèle, d'une part, et par les contraintes du logiciel de statistique, d'autre part, nous avons été contraints de séparer l'ensemble des instituts de sondages en deux groupes. Les résultats des régressions non linéaires apparaissent dans le tableau 6, disponible en annexe. Pour les deux sous-groupes d'instituts de sondages, nous retrouvons exactement le même type de résultat. Seuls le terme constant et les prix de marché relevés à la clôture des deux jours précédents extériorisent des coefficients statistiquement différents de zéro. Pour écarter un possible effet défavorable de la période de faible activité du marché, nous reconduisons l'expérience par deux fois avec une date de départ de l'échantillon au 1^{er} janvier 2007, puis au 1^{er} janvier 2008. Nos conclusions précédentes ne sont pas modifiées par ces nouveaux résultats.

5.3.1.3. *Les sondages sont-ils pour autant inutiles ?*

Les résultats de notre étude concernant l'élection présidentielle américaine de 2008 s'inscrivent dans l'exacte continuité des conclusions formulées par Forsythe, Nelson, Neumann et Wright (1992) dans leurs travaux menés sur l'élection présidentielle de 1988. Il est donc probable que, bien que le modèle développé par ces auteurs soit conceptuellement très séduisant, il ne permette pas d'expliquer les variations du prix de marché. Pour autant, pouvons-nous conclure que l'information contenue dans les sondages n'est d'aucune utilité pour prévoir le prix de marché ? Il apparaît évident que cette question ne peut être tranchée sans avoir mené de recherches complémentaires. Par ailleurs, après avoir mis en place et testé le modèle de Forsythe, Nelson, Neumann et Wright (1992), nous sommes sensibles aux arguments de Kou et Sobel (2001). Même si le modèle restitue parfaitement l'idée selon laquelle les prix de marché ne seraient influencés que par la « surprise » contenue dans le résultat d'un nouveau sondage, il reste très contraignant pour l'utilisation des données issues des sondages et ne permet pas d'examiner d'autres types de corrélations entre marché et sondages. En outre, nos tests sur la performance relative du marché ont mis en évidence que le sondage moyen pouvait représenter un bon

élément de prédiction de l'issue du scrutin. Pouvons-nous imaginer qu'il soit également un type d'information primordiale que les investisseurs prennent en compte au moment d'évaluer la valeur fondamentale de leurs contrats ? C'est ce que nous tentons de vérifier dans la section suivante.

5.3.2. Pouvons-nous construire un autre modèle ?

5.3.2.1. *Hypothèses*

Nous avons relevé lors de nos tests sur la performance relative du marché par rapport aux sondages (section 5.2.2.) que le sondage moyen quotidien pouvait s'avérer plus précis que le marché. Nous formulons ici l'hypothèse qu'il pourrait s'avérer tout aussi efficace pour reconstituer les prix de marché et tentons d'obtenir un modèle issu de régressions linéaires. Par ailleurs, nous testons également d'autres variables qui n'étaient pas considérées dans les recherches précédentes sur le sujet. Tout d'abord, nous recherchons un lien éventuel entre le prix de marché, que nous relevons cette fois-ci séparément pour chaque contrat et non plus sous la forme d'un écart normé entre les prix des deux contrats, et les résultats des sondages agrégés en moyenne quotidienne. Nous définissons deux séries de données : l'une contenant les moyennes relevées à leur date de publication, l'autre contenant les moyennes relevées le dernier jour de l'enquête terrain. Ces deux variables sont déclinées de façon temporelle en trois séries de données : une pour le jour considéré, une pour la veille de ce jour et une dernière pour l'avant-veille. Ensuite, nous testons une variable construite à partir des informations nouvelles et des événements de la journée. Si globalement les informations ont été favorables au candidat démocrate, cette variable prend la valeur 1. Si globalement les informations ont été favorables au candidat républicain, cette variable prend la valeur -1. Lorsque aucune information nouvelle n'est publiée ou lorsque les informations n'avantagent pas l'un ou l'autre des candidats, cette variable est nulle. Cette variable est construite à partir de la chronologie établie a posteriori par l'encyclopédie en ligne Wikipedia. Elle est volontairement subjective et vise à rechercher l'impact d'éventuels paramètres subjectifs et de mouvements d'opinion dans la construction des prix de marché. Nous proposons également une variable constituée par la moyenne des proportions d'indécis publiées dans les enquêtes d'opinion. Comme pour le sondage moyen, cette moyenne est calculée à deux reprises, soit à la date de publication du sondage, soit à la date du dernier jour de l'enquête terrain. De même, nous construisons une variable représentant un indice de confiance des intentions de vote exprimées dans les enquêtes d'opinion. L'idée sous-jacente est que les investisseurs pourraient appliquer une décote au prix de marché en fonction de l'intensité du soutien affiché à un candidat dans les enquêtes d'opinion. Nous calculons les valeurs de cette variable comme la proportion de

répondants manifestant un soutien fort au candidat. Le principal défaut de cette variable réside dans le peu d'observations disponibles : seules 55 enquêtes d'opinion sur les 642 de notre panel contiennent cette information. Lorsque l'information n'est pas disponible pour une journée donnée, nous conservons l'information disponible la plus récente. Enfin, nous testons l'impact des fluctuations des marchés financiers traditionnels (représentés par l'indice S&P500) sur les variations du prix de marché. A cet effet, nous disposons du niveau de clôture de l'indice sur les deux derniers jours, de la variation quotidienne de l'indice (en pourcentage) et de la moyenne hebdomadaire du niveau de l'indice.

5.3.2.2. *Les variables évidentes ne donnent pas de résultat*

Sur l'ensemble des variables que nous avons testées, la plupart se sont révélées non significatives. Nous nous attendions, en particulier, à relever une certaine influence de notre variable « effet de l'information » sur les prix de marché, et ce, malgré la subjectivité de la construction des données. Il n'en est rien, la régression montrant que l'on peut attribuer un coefficient nul à cette variable avec un niveau de confiance proche de 100%. Nous n'obtenons guère de meilleurs résultats avec nos variables mesurant l'indice de confiance des résultats ou avec l'une ou l'autre des variables représentant une certaine moyenne des résultats des sondages. Le tableau 7 de l'annexe montre les résultats des régressions que nous avons menées d'abord sur l'ensemble du panel et ensuite sur des périodes de temps plus réduites. Ces périodes de temps plus réduites nous permettent de tester la stabilité du modèle au cours du temps en ignorant les premiers mois d'ouverture du marché où l'activité est plus réduite et où, par conséquent, les variations de prix sont minimales. Si les modèles apparaissent sensiblement variables au cours du temps, la contribution du prix de marché relevé la veille est indiscutable. De toutes les variables étudiées séparément, c'est celle qui permet de prédire le plus précisément le prix de marché du lendemain. Les coefficients de la régression sont toujours statistiquement différents de zéro pour cette variable. Sur les premiers mois de l'ouverture du marché, nous remarquons également que le prix de marché relevé deux jours plus tôt est utilisé dans le modèle. Ceci s'explique, encore une fois, par le fait que la faible activité du marché peut conduire à un prix inchangé pendant deux, trois, voire même quatre jours consécutifs.

Enfin, il faut souligner l'impact, bien que limité à la seule période septembre-novembre 2008, de notre variable représentant la proportion d'indécis dans les répondants aux enquêtes d'opinion. Pour le modèle de cette période, le coefficient est positif, suggérant que plus la proportion d'indécis augmente, plus le contrat associé au candidat démocrate s'apprécie. Ce résultat est sans doute plus le fruit d'un hasard que d'une véritable corrélation. Si l'on peut justifier qu'en théorie une augmentation de la proportion d'indécis dans la population puisse être

interprétée comme une augmentation des réserves potentielles de voix pour un candidat en particulier (ou même pour les deux candidats), nous ne voyons pas pourquoi cet effet devrait être limité à une seule période. De plus, à mesure que l'on s'éloigne de la date de scrutin, la proportion d'indécis augmentant, cet effet devrait être démultiplié. Or, nous ne trouvons pas d'impact de cette variable sur le prix de marché au cours des autres périodes.

5.3.2.3. *Impact de l'économie sur le prix de marché*

Nous avons cependant été davantage surpris de découvrir que l'ajout de variables concernant le S&P500 permettait d'améliorer significativement le R^2 du modèle. Notons que ces variables apparaissent dans les modèles de chacune des périodes à l'exception de la période étudiée entre le 1^{er} janvier 2008 et le 7 novembre 2008. Même si les coefficients associés à ces variables sont parfois faibles (-0,059 pour le modèle sur la période débutant le 1^{er} janvier 2007), il convient de rappeler que les variables S&P500 que nous insérons dans le modèle sont exprimées en points (l'indice prenant des valeurs comprises entre 848 et 1565 points sur la période), à l'exception de la variation quotidienne de l'indice qui est exprimée en pourcentage. Par ailleurs, selon notre découpage temporel, les variables dérivées du S&P500 ont le plus d'influence sur le modèle établi entre le 1^{er} septembre 2008 et le 7 novembre 2008. A ce stade, le modèle semble donc nous amener à établir une relation de causalité entre les prix de l'IEM et l'instabilité boursière observée à cette période. Avant de formuler une telle conclusion, remarquons que les variables dérivées du S&P500 ont un meilleur effet lorsqu'elles sont associées. De plus, les coefficients associés à ces variables sont tous statistiquement significatifs au niveau de confiance de 95% et sont généralement de signes opposés. Lorsque nous étudions le prix de marché à une date t , le coefficient de la variable S&P à la date t est toujours positif, celui de la variable S&P500 à la date $t-1$ est toujours négatif, ainsi que celui de la variation journalière du niveau de l'indice. Nous comprenons ainsi que le modèle suggère qu'une baisse de l'indice est associée à une augmentation du prix du contrat relatif au candidat démocrate (et inversement pour le candidat républicain). Nous voyons deux explications à ce phénomène. La première consiste à croire que l'IEM est utilisé par les investisseurs comme marché de couverture : lorsque l'indice S&P clôture en repli, le contrat du candidat démocrate s'apprécie, ce qui permet de couvrir les pertes encourues sur les marchés conventionnels. Un tel mécanisme nous paraît difficilement justifiable, notamment parce que le montant total des investissements de chaque individu est limité à \$500. Ensuite, si un tel mécanisme de couverture était mis en place, nous devrions retrouver les variables dérivées de l'indice S&P500 dans les modèles de chaque période ou, si ce mécanisme ne s'était mis en place que tardivement au cours du mois de septembre 2008, dans le modèle de la dernière période (octobre-novembre 2008). La seconde explication interprète les variables du

S&P500 non comme des indicateurs des mouvements boursiers mais comme l'expression du climat économique du pays. Les baisses du niveau de l'indice sont interprétées comme des signes de perte de confiance des investisseurs dans la poursuite de la croissance économique et, par conséquent, l'expression d'une crainte d'un ralentissement économique. Le lien entre la sphère boursière et la sphère économique est d'ailleurs souvent relayé par les médias. Le modèle indiquerait donc que de mauvaises perspectives économiques entraînent une appréciation du contrat du candidat démocrate. Cette explication est d'autant plus plausible que le mois de septembre 2008 a vu se développer de nombreuses critiques du candidat républicain, le décrivant notamment incompetent en matière économique. L'idée selon laquelle une révision à la baisse des anticipations économiques profiterait au candidat démocrate est donc tout à fait plausible.

5.3.2.4. *L'impact des sondages n'est pas mesurable*

A la lumière de nos résultats, il semblerait que nous devions rejeter notre hypothèse de départ. L'agrégation des résultats des sondages en moyenne quotidienne ne permet pas d'expliquer les variations des prix de marché, pas plus que les séries temporelles des résultats publiés par chaque institut de sondage ne le permettrait, tout au moins dans le cadre du modèle de Forsythe, Nelson, Neumann et Wright (1992). Nous mettons cependant en évidence l'impact d'autres variables, et notamment de variables de type économique. Notre modèle principal de régression ne nous permet toutefois pas d'expliquer l'intégralité des variations des prix de marché. Il convient enfin de souligner que le principal facteur explicatif du modèle reste le prix en vigueur à la clôture de la veille. Tout se passe donc comme si le prix de marché était une martingale : la meilleure prédiction du prix futur, ou l'espérance du prix futur, n'est autre que le prix actuel. Cette constatation s'inscrit dans la droite ligne des hypothèses de marche aléatoire (*Random Walk*) du prix de marché formulées par Berg, Nelson et Rietz (2003). Les prix du *Vote Share Market* seraient donc soumis à un processus aléatoire. Cette conclusion reste néanmoins frustrante en pratique : puisque les échanges sur le marché fixent les prix et que ces échanges sont décidés par des investisseurs qui évaluent la valeur des contrats en fonction du set informationnel qu'ils ont à leur disposition, comment est-il possible que les sondages, qui sont l'information relative au scrutin la plus immédiatement disponible, ne leur soient d'aucune utilité dans ce processus d'évaluation ? Dans la section suivante, nous tentons d'évaluer l'intérêt pour les investisseurs de l'information contenue dans les sondages en développant une théorie plus sélective.

5.4. Certains Etats ont-ils plus d'influence que d'autres sur les prix du marché ?

Nous concluons notre travail de recherche en nous proposant de tester l'impact des résultats des sondages de chaque Etat considéré individuellement. Nous cherchons ainsi à mettre en évidence une éventuelle influence des *swing states* sur les prix de marché.

5.4.1. Les Swing States : la clé du prix de marché ?

Nous formulons maintenant l'hypothèse que tous les sondages ne constituent pas nécessairement une information de valeur pour l'investisseur. Suivant la logique établie par Forsythe, Nelson, Neumann et Wright (1992), nous supposons que seules les informations non anticipées par le marché, les « surprises », permettent d'expliquer les mouvements de prix sur l'PIEM. A ce titre, certains résultats de sondages locaux ne sont pas des « surprises » au sens où la victoire de l'un des deux principaux candidats est clairement anticipée dans quelques Etats. Comme nous le soulignons précédemment (section 4.2.3.), l'Idaho et l'Oklahoma n'ont plus jamais voté démocrate depuis l'élection de 1964 alors que d'autres Etats, comme Rhode Island et Hawaï, sont réputés pour voter systématiquement pour le candidat démocrate. Ainsi, un résultat d'enquête d'opinion annonçant une victoire du candidat démocrate à Rhode Island ou du candidat républicain dans l'Oklahoma ne devrait avoir aucun impact sur les prix de marché. En revanche, un sondage annonçant une possible victoire du candidat républicain à Hawaï devrait avoir un impact significatif sur les résultats. De la même façon, en raison de l'utilisation du système *Winner-Take-All* dans l'attribution des représentants du collège électoral dans chaque Etat, les résultats des sondages menés dans les Etats où le vote sera, a priori, le plus indécis, devraient avoir un impact significatif sur les prix de marché. Ces Etats qui peuvent faire basculer le résultat d'une élection présidentielle sont appelés *swing states* car ils sont susceptibles de changer de camp politique d'une élection à l'autre et, de fait, ont une grande influence sur le résultat de l'élection.

Pour vérifier cette théorie, nous utilisons notre base de 1 458 sondages locaux et définissons, pour chaque Etat, plusieurs variables clés. Tout d'abord, nous construisons une variable à partir des pourcentages d'intentions de vote exprimées par les répondants, pour chacun des deux principaux candidats, relevés à la date correspondant au dernier jour de l'enquête terrain. Cette variable a une valeur nulle lorsque aucune enquête d'opinion ne s'est terminée le jour considéré. Ensuite, nous définissons une variable qui garde en mémoire le dernier résultat du sondage publié pour un Etat donné. Cette variable est modifiée par toute valeur non nulle de notre première variable. Enfin, nous disposons d'une variable qui retient l'écart d'intentions de

vote exprimées pour chaque candidat entre deux enquêtes d'opinion successives. Pour chacun des 50 Etats américains, nous obtenons donc trois prédictors potentiels du prix de marché, soit un total de 150 variables disponibles. Compte tenu de la quantité de variables à tester et du peu de sondages locaux publiés avant l'année 2008, nous limitons notre recherche à la période courant entre le 1^{er} janvier 2008 et le 7 novembre 2008. Nous vérifions également la stabilité du modèle sur une période plus réduite, du 1^{er} août 2008 au 7 novembre 2008. Dans un second temps, nous introduisons une nouvelle variable, égale au prix relevé à la clôture du *Winner-Take-All Market* la veille du jour considéré dans le modèle, afin de contrôler l'impact du prix de marché en temps que variable explicative, tel que nous l'avions relevé au niveau fédéral. Notre recherche porte en premier lieu sur les prix du contrat DEM08_WTA.

5.4.2. Des résultats contraires aux hypothèses sur le contrat démocrate

Les modèles linéaires que nous présentons dans le tableau 8 sont obtenus par régressions successives du prix de marché sur un ensemble de variables. Avec plus de 150 variables disponibles, il est certain que nous pouvons trouver une combinaison linéaire de ces 150 variables qui réplique presque exactement le prix de marché. Afin d'obtenir un modèle le plus explicatif possible, nous retirons systématiquement toutes les variables non significatives au niveau de confiance 85% dans un premier temps, puis 95% pour toutes les étapes ultérieures. Nous procédons à une nouvelle régression sur les variables restantes et excluons les variables non significatives jusqu'à obtenir des coefficients qui soient tous statistiquement différents de zéro. Le modèle finalement obtenu est celui que nous présentons dans le tableau 8.

Bien que le premier modèle, représentant les prix de marché sur la période du 1^{er} janvier 2008 au 7 novembre 2008, contienne dix-sept variables, il possède le R^2 le plus faible de nos quatre modèles. Sur ces dix-sept variables, seules trois correspondent à des variations entre deux sondages consécutifs, les quatorze autres étant constituées des derniers résultats de sondage disponibles. Sur les quatorze variables utilisant les résultats des sondages locaux, huit proviennent d'Etats où le candidat républicain s'est finalement imposé. Il est intéressant de noter que le premier modèle (et aucun des trois autres modèles que nous présenterons par la suite) n'utilise pas les résultats des sondages qui seront publiés le lendemain ou le surlendemain et qui constituaient notre première catégorie de variables. Si tel avait été le cas, cela aurait suggéré que le marché était capable d'anticiper la publication des résultats des sondages. Même si Forsythe, Nelson, Neumann et Wright (1992) affirmaient que le marché était capable de prévoir les variations des résultats des enquêtes d'opinion, nous ne retrouvons pas une telle caractéristique au niveau des sondages locaux. Les signes négatifs de certains coefficients sont par ailleurs très troublants. Ils suggèrent, par exemple, qu'une hausse des intentions de vote pour le candidat

démocrate dans le Massachusetts a un impact négatif sur le prix de marché. Il en irait de même pour les résultats en provenance de la Virginie, alors que les mêmes variations relevées dans le Tennessee auraient un effet positif sur le prix du contrat démocrate. De même, selon le modèle, les intentions de vote des électeurs du Mississippi, du Tennessee, du Texas et du Wyoming influeraient négativement sur le prix de marché. Le vote de ces électeurs est-il si défavorable ? Est-il, par exemple, de nature à pousser l'ensemble des électeurs des autres Etats à modifier leur vote pour ne pas voter de la même façon que ces Etats ? Il est possible qu'une hausse des intentions de vote en faveur du candidat démocrate dans ces Etats soit le résultat d'annonces ou de promesses conservatrices de la part du candidat démocrate, qui ne plaisent pas à une partie de son électorat et soient de nature à réduire ses chances de remporter l'élection au niveau fédéral.

Le modèle, tel qu'il ressort de la régression linéaire, anéantit également notre théorie de l'impact des *swing states* sur le prix de marché. Bien au contraire, à l'exception du Montana, tous les autres résultats de sondages utilisés par le modèle correspondent à des Etats où un des deux candidats s'est imposé avec une large majorité (plus de quatre points d'écart). Le candidat républicain a en effet remporté les Etats du Texas, du Mississippi, du Tennessee, du Kansas et du Wyoming avec plus de 10 points d'avance sur son concurrent. Dans l'Oregon, le Nevada et le New Hampshire, le candidat démocrate remporte l'élection avec plus de 10 points d'avance, et son avance atteint 26 points dans le Massachusetts. Si l'on en croit le modèle, pour estimer la valeur du contrat à terme du candidat démocrate sur le *Winner-Take-All Market*, il faudrait non pas scruter avec attention les Etats susceptibles de changer de camp politique mais s'intéresser aux Etats qui voteront massivement pour un des deux candidats et analyser les variations des intentions de vote dans ces Etats.

Nous ne pouvons pas non plus écarter l'hypothèse selon laquelle le modèle ne restituerait plus ou moins fidèlement les variations du prix de marché que par le fruit du hasard, les variables utilisées étant suffisamment volatiles pour qu'une combinaison linéaire de certaines d'entre elles puisse reproduire à peu près toute série de données. Nous serions probablement moins sévères à l'égard du modèle s'il avait résisté au raccourcissement de la période d'observation que nous imposons ensuite. Le second modèle que nous obtenons ne conserve que sept variables sur les dix-sept du premier modèle, les autres étant devenues non significatives. Sur les sept variables restantes, une seule provient d'un Etat qui sera finalement remporté par le candidat démocrate, tous les autres Etats du modèle votant républicain. Par ailleurs, les coefficients des variables qui subsistent ont des valeurs relativement éloignées de celles relevées dans le premier modèle, et un des coefficients change même de signe. Il faudrait désormais croire que l'effet négatif des intentions de vote des électeurs du Tennessee devient positif à l'approche du scrutin. Si le modèle

reproduit très fidèlement les variations du prix de marché (avec un R^2 de 0,926), sa justification reste subjective.

Toutefois, lorsque nous donnons au modèle la possibilité d'intégrer les données de prix passées, cette variable est systématiquement présente dans le modèle avec un coefficient statistiquement significatif, quelle que soit la période de temps considérée. Par ailleurs, bien que certaines variables présentent dans le premier modèle persistent, tels les résultats des sondages du Kansas, du Nevada, du New Hampshire et du Texas, ils n'apportent pas un supplément de précision du modèle très sensible. Entre le 1^{er} août 2008 et le 7 novembre 2008 par exemple, le modèle à une variable $M_t = \alpha + \beta \cdot M_{t-1} + \varepsilon_t$ permet de reproduire plus de 90% des variations du prix de marché du contrat DEM08_WTA. Nos résultats restent donc difficiles à interpréter : pouvons-nous raisonnablement supposer que le premier modèle obtenu n'est pas le fruit du hasard ? Peut-on prévoir les prix de marché en observant les résultats des sondages locaux dans les Etats fortement marqués politiquement ? Pour s'en assurer, nous reconduisons la même recherche, mais sur les prix du contrat républicain.

5.4.3. Vérification des résultats par analyse du contrat REP08_WTA

Par bien des égards, les résultats de la régression linéaire menée sur le prix du contrat REP08_WTA, que nous proposons au lecteur en annexe (Tableau 9), diffèrent de nos premiers résultats. Il nous faut tout d'abord vingt-deux variables pour constituer le modèle sur l'année 2008, dix-huit variables représentant les derniers résultats connus dans chaque Etat et quatre représentant les variations des intentions de vote pour le candidat républicain entre les deux dernières enquêtes publiées. Nous retrouvons des coefficients positifs et des coefficients négatifs dans le modèle. Cette fois-ci, notre interprétation possible du signe négatif semble être invalidée : quelle homogénéité peut-on trouver entre les soutiens du candidat républicain de l'Idaho et du Maine ou du Minnesota ? Bien que ces trois variables soient présentes dans le modèle avec un coefficient négatif, le candidat républicain s'est imposé dans le premier Etat avec une avance de plus de trente points alors qu'il a concédé plus de 10 points à son adversaire démocrate dans les deux autres Etats. Par ailleurs, si le grand nombre de variables incorporées au modèle permet de restituer fidèlement les variations du prix de marché (avec un R^2 de 0,931), il n'en reste pas moins que l'erreur d'estimation est directement liée au nombre de variables du modèle. Chaque résultat de sondage est en effet sujet à une marge d'erreur (en moyenne 4% pour notre échantillon), cumuler 22 variables implique donc de prendre en compte ce facteur d'incertitude. Lorsque nous réduisons la période d'observation, le modèle du prix du contrat républicain s'avère très instable. Ce second modèle ne conserve qu'une seule variable du premier et contient neuf variables. Enfin,

le fait de donner au modèle la possibilité d'incorporer une variable de prix historique écarte toutes les autres variables. Pour prévoir le prix de marché du contrat républicain il semble donc beaucoup plus efficace d'observer le prix de clôture relevé la veille sur le marché.

5.4.4. Conclusion

Contrairement aux hypothèses que nous avons formulées a priori, nos recherches ne permettent pas de mettre en évidence un impact des résultats des sondages publiés dans les *swing states* sur les prix des contrats du *Winner-Take-All Market*. D'autre part les variables explicatives que nous considérons dans notre modèle de prix du contrat démocrate ne semblent avoir répliqué les variations du prix de marché que par le simple hasard. Le fait que nous ne trouvions pas de relation similaire dans l'analyse du modèle de prix du contrat républicain soutient cette conclusion.

Après avoir analysé à la fois l'impact des sondages nationaux et des sondages locaux sur les prix de l'IEM, nous pouvons conclure que les résultats de ces enquêtes n'ont pas d'impact directement mesurable sur le marché. Ces résultats concordent avec les travaux de Forsythe, Nelson, Neumann et Wright (1992) et Bohm et Sonnegård (1999). En revanche, rien ne nous permet d'affirmer que les informations contenues dans les sondages n'influencent pas les investisseurs actifs sur les marchés de l'IEM. Le propre du marché est en effet de mettre en place un processus de collecte et d'agrégation de l'information pour parvenir à un prix d'équilibre. Il est donc possible que les résultats des sondages soient une des dizaines ou des centaines de variables qui influencent quotidiennement le prix de marché. L'impact d'une seule variable (les résultats des sondages) serait donc extrêmement difficile à mesurer, de même que l'impact de toute autre variable prise séparément.

Toutefois, bien que nous ne trouvions pas de relation certaine entre les résultats des sondages dans chaque Etat et le prix des contrats démocrate et républicain du *Winner-Take-All Market*, le caractère hautement explicatif de la variable historique des prix est un argument en faveur de l'hypothèse semi-forte de l'efficience des marchés pour l'IEM. Nous montrons en effet qu'un certain niveau d'information publique ne permet pas de prévoir les prix futurs. De plus, il semble tout aussi difficile d'établir la valeur future d'un contrat. La meilleure prévision du prix de clôture du lendemain reste apparemment le prix du jour, tel que Berg, Nelson et Rietz (2003) le suggèrent.

6. Conclusion

Notre recherche avait pour principaux objectifs de tester la performance des marchés de l'IEM comme éléments de prédiction du résultat de l'élection présidentielle américaine de 2008, à la fois en termes absolu et relatif, et d'étudier l'impact éventuel des résultats des sondages sur les prix de marché. Elle se proposait notamment d'étudier cet impact sur deux échantillons de résultats de sondages, l'un établi au niveau fédéral et l'autre constitué à partir des résultats des enquêtes d'opinion relevés dans chaque Etat de la fédération.

Nous avons mis en évidence la précision du prix de marché à court terme, à moyen terme et à long terme. Les cotations observées sur le *Vote Share Market* et sur le *Winner-Take-All Market* donnent des estimations très proches du résultat de l'élection, et ce plusieurs mois avant la date du scrutin. Quiconque aurait voulu obtenir une prévision fiable du résultat de l'élection présidentielle à un moment de la campagne électorale aurait pu se fier aux marchés de l'IEM. A bien des égards, nous pourrions également conseiller à cette personne de privilégier le marché plutôt que les enquêtes d'opinion pour obtenir une prévision la plus précise possible du résultat de l'élection. Nous avons cependant observé que, pour ce qui concerne l'élection de 2008, l'agrégation des résultats des sondages publiés chaque jour donne en général une prévision plus précise que le marché, sur les quatre dernières semaines de la campagne électorale.

Nous ne sommes pas parvenus à déterminer une relation de causalité entre les résultats des enquêtes d'opinion et les prix de marché, que nous ayons relevé ceux-ci sur le *Vote Share Market* ou sur le *Winner-Take-All Market*. Nous observons cependant sur le premier marché l'impact probable sur le prix de marché d'une variable reflétant les perspectives de croissance économique. Si nous ne pouvons prouver formellement la relation entre cette variable et le prix de marché, il est fortement probable que les informations économiques soient de nature à modifier la valorisation des contrats par les investisseurs.

Enfin, nous ne pouvons pas évaluer l'impact des événements conjoncturels dans notre analyse : le choix des tickets par les deux grands partis a-t-il eu un impact sur les cotations du marché ? Le déroulement de la campagne a-t-il provoqué des déviations importantes de la valeur fondamentale sur les marchés de l'IEM ? Ces thématiques font partie d'un nombre important de sujets de recherche qui s'inscrivent dans la continuité de nos travaux et qui n'ont pas encore trouvés de réponses. L'un des sujets les plus passionnants à nos yeux est l'étude de l'impact des publications d'information sur les cotations du marché. Nous n'avons encore que trop peu d'informations sur la façon dont les marchés de l'IEM réagissent à certaines informations clés de la campagne. Nous ne savons pas non plus quels sont les types d'informations qui sont de nature

à modifier le comportement des investisseurs et celles qui sont sans effet. Nous ignorons enfin la rapidité d'incorporation de ces informations au prix de marché et les facteurs qui influencent cette rapidité. Ces questions sont fondamentales dans la compréhension du fonctionnement des marchés prédictifs tels que ceux de l'Iowa Electronic Markets. Y apporter des réponses requière néanmoins de disposer d'une qualité d'information largement supérieure à celle dont nous disposons pour cette étude. Elle implique, à notre avis, de recueillir l'ensemble des prix de transaction observés sur les marchés de l'IEM et, possiblement, l'état périodique du carnet d'ordres aux moments clés de publication des nouvelles informations. Nous ne sommes par ailleurs pas certains que l'Université de l'Iowa conserve indéfiniment une telle quantité d'informations.

Le potentiel des marchés prédictifs en général, et des marchés sur les élections politiques américaines en particulier, reste donc largement sous-exploité. L'éventail de ses utilisations pratiques est tout aussi méconnu. Si beaucoup pensent que l'échange des contrats financiers sur les candidats à l'élection présidentielle américaine ne revêt que peu d'intérêt, nous pensons au contraire qu'il s'agit là d'une forme d'utilisation ludique qui présage d'une utilisation étendue à l'ensemble des secteurs économiques. Des expériences ont ainsi montré que lorsque de tels marchés sont établis à l'intérieur de l'entreprise, ils s'avèrent extrêmement performants lorsqu'il s'agit de prévoir la réalisation d'événements futurs en lien avec l'entreprise. On peut ainsi imaginer un champ d'application considérable aux marchés prédictifs et la popularité de l'Iowa Electronic Markets ne fait qu'augurer le succès de ces futurs outils de prédiction.

Bibliographie

Berg, J., Forsythe, R., Nelson, F., Rietz, T., 2000. Results from a Dozen Years of Election Futures Markets Research. Université de l'Iowa, Document de travail.

Berg, J., Forsythe, R., Rietz, T., 1997. The Iowa Electronic Market, in : Passon, D., Wood, D. (Eds.), Blackwell Encyclopedic Dictionary of Finance. Royaume-Uni, Oxford, pp. 219-222.

Berg, J., Nelson, F., Rietz, T., 2003. Accuracy and Forecast Standard Error of Prediction Markets. Université de l'Iowa, Document de travail.

Berg, J., Nelson, F., Rietz, T., 2008. Prediction Market Accuracy in The Long Run. International Journal of Forecasting 24, pp. 283-298.

Berg, J., Rietz, T., 2002. Longshots, Overconfidence and Efficiency on the Iowa Electronic Market. Université de l'Iowa, Document de travail.

Berg, J., Rietz, T., 2006. The Iowa Electronic Markets: Stylized Facts and Open Issues, in : Hahn, R., Tetlock, P. (Eds.), Information Markets: A New Way of Making Decisions. Etats-Unis, Washington, D.C., pp. 142-169.

Bohm, P., Sonnegård, J., 1999. Political Stock Markets and Unreliable Polls. Scandinavian Journal of Economics, 101, pp. 205-222.

Boyle, G., Videbeck, S., 2006. Want to Predict The Future: Ask the Market! Policy, Vol. 22, No. 2, pp. 39-42.

Crespi, I., 1988. Pre-Election Polling: Sources of Accuracy and Error. Russel Sage Foundation.

Erikson, R., Wlezien, C., 2008. Are Political Markets Really Superior to Polls As Election Predictors? Public Opinion Quarterly, Vol. 72, No. 2, pp. 190-215.

Fair, R., 1978. The Effect of Economic Events on Votes for President. The Review of Economics and Statistics, Vol. LX, No. 2, pp. 159-173.

Financial Executive. Prediction Markets: Crystal Ball For Finance? Mai 2004, pp. 24-26.

Forsythe, R., Nelson, F., Neumann, G., Wright, J., 1992. Anatomy of an Experimental Political Stock Market. *The American Economic Review*, Vol. 82, No. 5, pp. 1142-1161.

Forsythe, R., Rietz, T., Ross, T., 1998. Wishes, Expectations and Actions: Price Formation in Election Stock Markets. *Journal of Economic Behavior and Organization*, Vol. 39, pp. 83-100.

Gelman, A., King, G., 1993. Why Are American Presidential Election Campaign Polls so Variable When Votes Are so Predictable? *British Journal of Political Science*, Vol. 23, No. 4, pp. 409-451.

Gomme, P., 2003. Iowa Electronic Markets. Federal Reserve Bank of Cleveland. États-Unis, Cleveland.

Investment Dealer's Digest. With Two Months To Go, Iowa Futures Market Says Gore Will Win By a Slim Margin. 11 septembre 2000, p. 17.

Kou, S., Sobel, M., 2004. Forecasting the Vote: An Analytical Comparison of Election Markets and Public Opinion Polls. *Political Analysis*, Vol. 12, pp. 277-295.

Leigh, A., Wolfers, J., 2006. Competing Approaches to Forecasting Elections: Economic Models, Opinion Polls and Prediction Markets. *The Economic Record*, Vol. 82, No. 258, pp. 325-340.

Levmore, S., 2003. Simply Efficient Markets and the Role of Regulation: Lessons from the Iowa Electronic Markets and the Hollywood Stock Exchange. *The Journal of Corporation Law*, July 2003, pp. 589-606.

Lockerbie, B., 2004. A Look to the Future: Forecasting the 2004 Presidential Election. *PS: Political Science and Politics*, Vol. 37, No. 4, pp. 741-743.

Lockerbie, B., 2005. After the Vote: Evaluating a Prospective Forecasting Model of Presidential Elections. *PS: Political Science and Politics*, Vol. 38, No. 1, pp. 39-40.

Oliven, K., Rietz, T., 2004. Suckers Are Born but Markets Are Made: Individual Rationality, Arbitrage, and Market Efficiency on an Electronic Futures Market. *Management Science*, Vol. 50, No. 3, pp. 336-351.

Rhode, P., Strumpf, K., 2004. Historical Presidential Betting Markets. *Journal of Economic Perspectives*, Vol. 18, No. 2, pp. 127-142.

Schonfeld, E., 2006. The Wisdom of the Corporate Crowd. *Business 2.0*, Vol. 7, No. 8, pp. 47-49.

Stix, G., 2008. When Markets Beat the Polls. *Scientific American*, Vol. 298, No. 3, pp. 38-45.

Surowiecki, J., 2004. *The Wisdom of Crowds*. Royaume-Uni, Londres.

The Economist. A Market For Votes. 11 avril 2000, édition en ligne.

The Economist. Guessing Games. 20 novembre 2004, p. 80.

The Wall Street Journal. Iowa Market Takes Stock of Presidential Candidates. 28 août 1995.

UI News Services (*Service de presse de l'Université de l'Iowa*). IEM within Less Than Half Percentage Point in Presidential Race Prediction. 24 novembre 2008.

Wikipedia.org. Encyclopédie libre en ligne.

Wolfers, J., Zitewitz, J., 2004. Prediction Markets. *Journal of Economic Perspectives*, Vol. 18, No. 2, pp. 107-126.

Annexes

Ecran de trading de l'IEM

Trader: mmm		NewActivity		Iowa Electronic Markets		
PracticeS: 996.88700		Refresh		Prac VS		
5/24/2009 1:58:39 AM						
Contract	BestBid	BestAsk	LastPrice	QuantityHeld	YourBids	YourAsks
V.TRU	0.555*	----	0.575	<u>2</u>	<u>1</u>	<u>0</u>
V.DEW	----	----	0.430	<u>0</u>	<u>0</u>	<u>0</u>

-Market Orders-	▼	Qty	<input type="text"/>	Confirm?	<input checked="" type="checkbox"/>	Market Order
-Limit Orders-	▼	Price	<input type="text"/>	Qty	<input type="text"/>	Expire 05/26/09 11:59PM
				Confirm?	<input checked="" type="checkbox"/>	Limit Order

Prac_VS	▼	Go To Market	Market Info	My Account	Help	Logout
---------	---	--------------	-------------	------------	------	--------

L'écran ci-dessus reproduit les informations à disposition des investisseurs. Pour chacun des contrats du marché, l'investisseur connaît les meilleures cotations *bid* et *ask*, le prix réglé lors de la dernière transaction sur le marché ainsi que la quantité de chaque contrat qu'il possède en portefeuille et les ordres à cours limités qu'il a envoyés à l'IEM et restent en attente d'exécution dans le carnet d'ordres. L'astérisque placé derrière une cotation signifie que l'investisseur dispose de la meilleure cotation dans cette partie du carnet d'ordres.

Depuis cet écran, l'investisseur peut également envoyer à l'IEM des ordres au mieux ou des ordres à cours limité pour prendre position sur les contrats du marché.

Graphique 6 : Comparaison des écarts normés des principaux résultats de sondages, du prix de marché et du résultat de l'élection

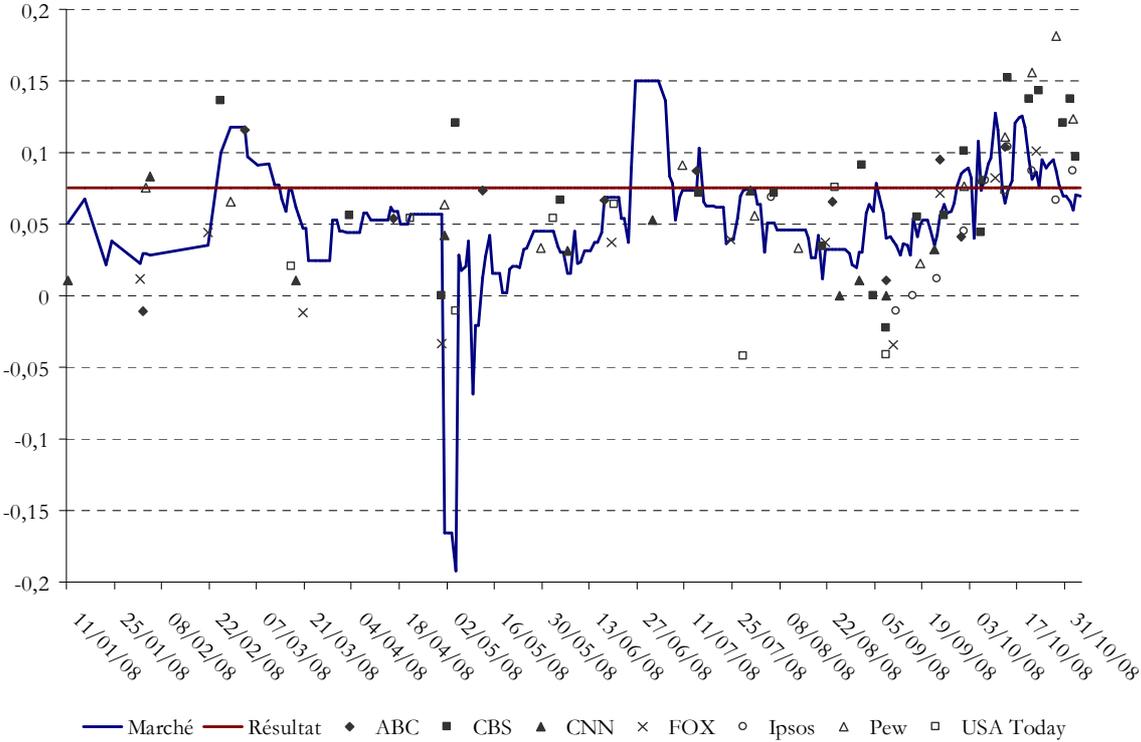


Tableau 1 : Présentation de l'échantillon de sondages nationaux

	Code	Sondages ponctuels	Trackers quotidiens	<i>Total</i>
ABC	A	11		11
CBS News	B	21		21
CNN	C	14		14
Cook Political Reports	O	6		6
Democracy Corps	S	8		8
DIAGEO	D	9	58	67
FOX News	F	13		13
GALLUP	G		229	229
GFK	K	3		3
IPSOS	I	10		10
LA Times	L	8		8
MARIST	M	4		4
NBC News	N	11		11
NEWSWEEK	W	9		9
PEW RESEARCH CENTER	P	13		13
QUINNIPIAC UNIVERSITY POLL	Q	9		9
RASMUSSEN REPORTS	R		147	147
TIME	T	3		3
USA TODAY	U	9		9
YAHOO NEWS	Y	6		6
ZOGBY	Z	12	29	41
<i>Total</i>		179	463	642
<i>Fréquence</i>		<i>27,9%</i>	<i>72,1%</i>	<i>100%</i>

Tableau 2 : Présentation de l'échantillon de sondages locaux

Etat	Nombre de Sondages	Fréquence	Nombre de représentants
Alabama	24	1,6%	9
Alaska	20	1,4%	3
Arizona	24	1,6%	10
Arkansas	30	2,1%	6
California	30	2,1%	55
Colorado	53	3,6%	9
Connecticut	14	1,0%	7
Delaware	8	0,5%	3
Florida	89	6,1%	27
Georgia	40	2,7%	15
Hawaï	2	0,1%	4
Idaho	4	0,3%	4
Illinois	14	1,0%	21
Indiana	34	2,3%	11
Iowa	37	2,5%	7
Kansas	18	1,2%	6
Kentucky	24	1,6%	8
Louisiana	9	0,6%	9
Maine	19	1,3%	4
Maryland	6	0,4%	10
Massachussets	21	1,4%	12
Michigan	51	3,5%	17
Minnesota	50	3,4%	10
Mississippi	15	1,0%	6
Missouri	47	3,2%	11
Montana	19	1,3%	3
Nebraska	7	0,5%	5
Nevada	38	2,6%	5
New Hampshire	43	2,9%	4
New Jersey	42	2,9%	15
New Mexico	28	1,9%	5
New York	36	2,5%	31
North Carolina	73	5,0%	15
North Dakota	11	0,8%	3
Ohio	97	6,7%	20
Oklahoma	9	0,6%	7
Oregon	32	2,2%	7
Pennsylvania	90	6,2%	21
Rhode Island	5	0,3%	4
South Carolina	13	0,9%	8
South Dakota	10	0,7%	3
Tennessee	14	1,0%	11
Texas	17	1,2%	34
Utah	10	0,7%	5
Vermont	5	0,3%	3
Virginia	67	4,6%	13
Washington	37	2,5%	11
West Virginia	19	1,3%	5
Wisconsin	46	3,2%	10
Wyoming	7	0,5%	3
Washington, D.C.*	N/A	N/A	3
Total	1458	100,0%	538

**Bien que ne constituant pas un Etat, le district de Washington possède trois représentants au collège électoral.*

Tableau 3 : Résultats du test binomial

<u>Fréquence des sondages</u>		Durée restante avant l'élection (4 Novembre 2008)					
		<i>Ensemble du panel</i>	<i>Plus de 100 jours</i>	<i>De 62 à 100 jours</i>	<i>De 32 à 61 jours</i>	<i>De 6 à 31 jours</i>	<i>Moins de 5 jours</i>
Nombre de sondages		642	255	104	127	131	25
dont basés sur les <i>registered voters</i>		547	241	95	115	82	14
Fréquence de sondages <i>registered voters</i>		85,20%	94,51%	91,35%	90,55%	62,60%	56,00%

<u>Sondages à date de publication</u>		Durée restante avant l'élection (4 Novembre 2008)					
		<i>Ensemble du panel</i>	<i>Plus de 100 jours</i>	<i>De 62 à 100 jours</i>	<i>De 32 à 61 jours</i>	<i>De 6 à 31 jours</i>	<i>Moins de 5 jours</i>
Nombre de sondages		642	255	104	127	131	25
Meilleure précision du marché des sondages		452	173	75	104	83	17
		190	82	29	23	48	8
% de meilleure précision du marché		70,40%	67,84%	72,12%	81,89%	63,36%	68,00%
Test binomial (valeur de p)		0,000	0,000	0,000	0,000	0,003	0,108

<u>Sondages au dernier jour de l'enquête</u>		Durée restante avant l'élection (4 Novembre 2008)					
		<i>Ensemble du panel</i>	<i>Plus de 100 jours</i>	<i>De 62 à 100 jours</i>	<i>De 32 à 61 jours</i>	<i>De 6 à 31 jours</i>	<i>Moins de 5 jours</i>
Nombre de sondages		642	255	104	127	131	25
Meilleure précision du marché des sondages		435	169	72	100	77	17
		207	86	32	27	54	8
% de meilleure précision du marché		67,76%	66,27%	69,23%	78,74%	58,78%	68,00%
Test binomial (valeur de p)		0,000	0,000	0,000	0,000	0,054	0,108

<u>Sondages agrégés en moyenne quotidienne</u>		Durée restante avant l'élection (4 Novembre 2008)					
		<i>Ensemble du panel</i>	<i>Plus de 100 jours</i>	<i>De 62 à 100 jours</i>	<i>De 32 à 61 jours</i>	<i>De 6 à 31 jours</i>	<i>Moins de 5 jours</i>
Nombre de sondages		260	159	39	30	27	5
Meilleure précision du marché des sondages		172	105	30	26	10	1
		88	54	9	4	17	4
% de meilleure précision du marché		66,15%	66,04%	76,92%	86,67%	37,04%	20,00%
Test binomial (valeur de p)		0,000	0,000	0,001	0,000	0,248	0,375

Tableau 4 : Activité historique des marchés de l'IEM détaillée par périodes

Vote Share Markets[†]					
	1992	1996	2000	2004	2008
<i>Ouverture du marché</i>	09/01/1992	25/10/1994	05/01/2000	21/02/2003	01/06/2006
<i>Clôture du marché</i>	08/11/1992	08/11/1996	10/11/2000	05/11/2004	07/11/2008
<i>Nombre de contrats</i>	7	2	3	16	2
Ensemble du panel					
<i>Volume échangé</i>	56 284	NC ¹	49 263	331 300	19 844
<i>Valeur des échanges</i>	\$12 389	NC ¹	\$14 816	\$47 455	\$10 012
Activité à plus de 100 jours du scrutin					
<i>Volume échangé</i>	45 637	NC ¹	16 068	294 167	6 807
<i>Valeur des échanges</i>	\$7 148	NC ¹	\$5 526	\$30 103	\$3 457
Activité entre 62 et 100 jours avant le scrutin					
<i>Volume échangé</i>	1 432	920	5 563	11 091	929
<i>Valeur des échanges</i>	\$716	\$480	\$1 648	\$4 319	\$494
Activité entre 32 et 61 jours avant le scrutin					
<i>Volume échangé</i>	4 181	1 001	8 148	5 290	5 023
<i>Valeur des échanges</i>	\$1 970	\$516	\$2 463	\$2 655	\$2 526
Activité entre 6 et 31 jours avant le scrutin					
<i>Volume échangé</i>	4 494	1 125	11 731	13 001	5 563
<i>Valeur des échanges</i>	\$2 283	\$585	\$2 999	\$6 503	\$2 766
Activité durant les 5 jours précédant le scrutin					
<i>Volume échangé</i>	540	732	5 058	6 472	699
<i>Valeur des échanges</i>	\$271	\$360	\$1 031	\$3 236	\$352
[†] Les données du Vote Share Market de 1988 ne sont pas présentées					
¹ Données non communiquées par l'IEM. Berg, Nelson et Rietz (2003) rapportent toutefois que 23 093 contrats ont été échangés pendant toute la durée d'ouverture du marché, pour un montant total de \$3,628.					
Winner-Take-All Markets					
	1992	1996	2000	2004	2008
<i>Ouverture du marché</i>	10/07/1992	15/11/1994	01/05/2000	01/06/2004	01/06/2006
<i>Clôture du marché</i>	04/11/1992	06/11/1996	10/11/2000	05/11/2004	07/11/2008
<i>Nombre de contrats</i>	3	4	3	4	2
Ensemble du panel					
<i>Volume échangé</i>	232 164	674 163	397 465	1 024 300	471 041
<i>Valeur des échanges</i>	\$53 547	141 258	\$152 588	\$327 385	\$216 188
Activité à plus de 100 jours du scrutin					
<i>Volume échangé</i>	13 883	298 549	27 646	33 091	141 680
<i>Valeur des échanges</i>	\$3 625	69 103	\$9 663	\$16 834	\$71 948
Activité entre 62 et 100 jours avant le scrutin					
<i>Volume échangé</i>	21 150	127 369	46 534	99 229	56 332
<i>Valeur des échanges</i>	\$5 663	\$28 895	\$16 163	\$52 599	\$28 538
Activité entre 32 et 61 jours avant le scrutin					
<i>Volume échangé</i>	40 654	102 830	77 476	236 317	108 334
<i>Valeur des échanges</i>	\$11 626	\$19 542	\$28 795	\$89 009	\$55 768
Activité entre 6 et 31 jours avant le scrutin					
<i>Volume échangé</i>	101 263	79 895	92 003	349 828	106 319
<i>Valeur des échanges</i>	\$23 744	\$15 108	\$41 847	\$97 469	\$44 666
Activité durant les 5 jours précédant le scrutin					
<i>Volume échangé</i>	55 163	65 179	51 934	210 111	51 820
<i>Valeur des échanges</i>	\$8 838	\$8 449	\$23 296	\$50 607	\$10 866

Tableau 5 : Résultats des coefficients du modèle de Forsythe, Nelson, Neumann et Wright (1992) pour les 4 Trackers

Trackers	1/06/06 - 7/11/08	12/03/08 - 7/11/08	8/06/08 - 7/11/08	11/09/08 - 7/11/08	7/10/08 - 7/11/08
α_0	0,250 E	0,021 E	- 0,147 E	- 0,581 E	0,042 E
α_1	0,520 (0,053)	0,726 (0,070)	0,692 (0,089)	0,406 (0,171)	0,957 (0,342)
α_2	0,186 (0,032)	0,035 (0,068)	0,170 (0,094)	0,103 (0,173)	- 0,479 E
Valeurs de β					
Diageo	0,028 (0,154)	0,051 (0,076)	0,089 (0,078)	0,000 (0,482)	0,204 (0,183)
Gallup	- 0,077 (0,282)	- 0,065 (0,110)	- 0,100 (0,076)	0,000 (0,160)	- 0,023 E
Rasmussen	0,283 (0,517)	0,278 (0,205)	0,361 (0,143)	0,107 (0,499)	- 0,500 (0,395)
Zogby/Reuters	- 0,030 (0,235)	- 0,440 (0,099)	- 0,082 (0,063)	0,102 (0,205)	0,292 (0,000)
R^2	0,489	0,687	0,785	0,700	0,712

Le tableau ci-dessus présente la valeur des coefficients de la régression. L'écart-type du coefficient est indiqué en-dessous de chaque coefficient, entre parenthèses.

La lettre "E" indique que la valeur de l'écart-type est supérieure à 1 000.

Tableau 6 : Résultats des coefficients du modèle de Forsythe, Nelson, Neumann et Wright (1992) pour l'ensemble de l'échantillon

Groupe 1				Groupe 2			
	Panel	1/01/07 - 7/11/08	1/01/08 - 7/11/08		Panel	1/01/07 - 7/11/08	1/01/08 - 7/11/08
α_0	0,170 (0,003)	0,041 (0,004)	0,130 (0,003)	α_0	0,170 (0,003)	0,039 (0,004)	0,110 (0,002)
α_1	0,519 (0,030)	0,115 (0,039)	0,741 (0,062)	α_1	0,522 (0,000)	0,140 (0,039)	0,773 (0,057)
α_2	0,185 (0,030)	0,098 (0,039)	0,021 (0,061)	α_2	0,186 (0,030)	0,119 (0,039)	0,031 (0,060)
Valeurs de β				Valeurs de β			
Diageo	0,384 E	0,580 E	0,509 E	Ipsos	0,735 E	0,616 E	0,504 E
Gallup	0,362 E	1,093 E	1,004 E	LA Times	0,677 E	0,700 E	0,660 E
Rasmussen	0,518 E	- 1,399 E	- 1,017 E	Marist	1,058 E	0,500 E	0,615 E
Zogby/Reuters	0,342 E	0,662 E	0,216 E	NBC	0,512 E	0,469 E	0,858 E
ABC News	0,372 E	0,613 E	0,328 E	Newsweek	0,531 E	0,544 E	0,319 E
CBS	0,350 E	0,583 E	0,343 E	Pew Research Center	0,812 E	0,789 E	0,678 E
CNN	0,386 E	0,457 E	0,426 E	Quinnipiac University	0,756 E	0,772 E	2,230 E
Cook Political Research	0,575 E	0,817 E	0,668 E	Time	0,567 E	0,868 E	0,647 E
Democracy Corps	0,552 E	0,785 E	0,531 E	USA Today	0,041 E	0,491 E	0,295 E
FOX News	0,382 E	0,769 E	0,487 E	Yahoo	0,358 E	0,358 E	0,553 E
GfK	0,406 E	0,855 E	0,355 E				
R²	0,491	0,086	0,717	R²	0,491	0,071	0,743

Le tableau ci-dessus présente la valeur des coefficients de la régression. L'écart-type du coefficient est indiqué en-dessous de chaque coefficient, entre parenthèses. La lettre "E" indique que la valeur de l'écart-type est supérieure à 1 000.

Tableau 7 : Résultats de la régression linéaire (sondages nationaux)

Modèle linéaire basé sur les sondages nationaux					
	1/06/06 - 7/11/08	1/01/07 - 7/11/08	1/01/08 - 7/11/08	1/09/08 - 7/11/08	1/10/08 - 7/11/08
M_{t-1}	0,898 (0,020)	0,709 (0,038)	0,659 (0,560)	0,604 (0,083)	0,774 (-0,112)
M_{t-2}	0,027 (0,016)	0,129 (0,038)	0,200 (0,560)		
$S\&P_t$	1,041 (0,000)	- 0,059 (0,000)		7,510 (0,000)	
$S\&P_{t-1}$	- 1,054 (0,000)			- 7,923 (0,000)	
$\Delta S\&P$	- 0,138 (0,134)			- 1,979 (0,214)	- 0,230 (0,039)
P_t^{moy}					
<i>Indécis</i>				0,306 (0,073)	
R ² ajusté	0,852	0,713	0,689	0,752	0,571

Le tableau ci-dessus présente la valeur des coefficients de la régression. L'écart-type du coefficient est indiqué en-dessous de chaque coefficient, entre parenthèses.

Tableau 8 : Résultats de la régression linéaire sur le prix du contrat démocrate DEM08_WTA (sondages locaux)

Modèles linéaires sans prix historiques				Modèle linéaire avec prix historiques			
	1/01/08 - 7/11/08		1/08/08 - 7/11/08		1/01/08 - 7/11/08		1/08/08 - 7/11/08
<i>Arizona(t-s)</i>	0,946 **	<i>Arizona(t-s)</i>	0,288 **	M_{t-1}	0,816 **	M_{t-1}	0,883 **
<i>Kansas(t-s)</i>	0,259 **	<i>Kansas(t-s)</i>	0,118 **	<i>Kansas(t-s)</i>	0,051 **	<i>Pennsylvanie(t-s)</i>	0,099 *
<i>Massachusetts(t-s)</i>	0,413 **	<i>Nevada(t-s)</i>	0,127 **	<i>Nevada(t-s)</i>	0,602 **		
<i>Mississippi(t-s)</i>	- 0,755 **	<i>Caroline du Sud (t-s)</i>	0,164 **	<i>New Hampshire(t-s)</i>	0,047 *		
<i>Montana(t-s)</i>	0,437 **	<i>Tennessee(t-s)</i>	0,235 **	<i>Texas(t-s)</i>	- 0,577 **		
<i>Nevada(t-s)</i>	1,672 **	<i>Texas(t-s)</i>	- 0,413 **				
<i>New Hampshire(t-s)</i>	0,234 **	<i>Wyoming(t-s)</i>	- 0,131 **				
<i>Ohio(t-s)</i>	0,134 **						
<i>Oregon(t-s)</i>	0,096 **						
<i>Caroline du Sud(t-s)</i>	0,542 **						
<i>Tennessee(t-s)</i>	- 1,092 **						
<i>Texas(t-s)</i>	- 1,692 **						
<i>Virginie(t-s)</i>	0,212 **						
<i>Wyoming(t-s)</i>	- 0,390 **						
$\Delta(\text{Massachusetts})$	- 0,070 **						
$\Delta(\text{Virginie})$	- 0,072 **						
$\Delta(\text{Tennessee})$	0,147 **						
R ² ajusté	0,859	R ² ajusté	0,926	R ² ajusté	0,912	R ² ajusté	0,912

* indique un coefficient statistiquement significatif au niveau de confiance 95%

** indique un coefficient statistiquement significatif au niveau de confiance 99%

* indique un coefficient statistiquement significatif au niveau de confiance 95%

** indique un coefficient statistiquement significatif au niveau de confiance 99%

Tableau 9 : Résultats de la régression linéaire sur le prix du contrat républicain REP08_WTA (sondages locaux)

Modèles linéaires sans prix historiques				Modèles linéaires avec prix historiques				
	1/01/08 - 7/11/08		1/08/08 - 7/11/08		1/01/08 - 7/11/08		1/08/08 - 7/11/08	
<i>Arizona(t-s)</i>	- 0,702 **	<i>Géorgie(t-s)</i>	4,769 **	M_{t-1}	0,987 **	M_{t-1}	0,998 **	
<i>Arkansas(t-s)</i>	- 0,380 **	<i>Nouveau Mexique(t-s)</i>	- 1,437 **					
<i>Californie(t-s)</i>	0,401 **	<i>Pennsylvanie(t-s)</i>	3,080 **					
<i>Géorgie(t-s)</i>	1,575 **	<i>Texas(t-s)</i>	- 2,670 **					
<i>Idaho(t-s)</i>	- 2,789 **	<i>Virginie occidentale(t-s)</i>	- 2,739 **					
<i>Iowa(t-s)</i>	0,051 *	$\Delta(\text{Kansas})$	- 0,068 **					
<i>Kentucky(t-s)</i>	0,238 **	$\Delta(\text{Nebraska})$	3,285 **					
<i>Louisiane(t-s)</i>	1,276 **	$\Delta(\text{Oregon})$	- 0,067 **					
<i>Maine(t-s)</i>	- 0,963 **	$\Delta(\text{Pennsylvanie})$	- 0,077 **					
<i>Minnesota(t-s)</i>	- 0,062 **							
<i>Montana(t-s)</i>	1,826 **							
<i>Nebraska(t-s)</i>	- 1,196 **							
<i>New Jersey(t-s)</i>	0,120 **							
<i>Caroline du Nord(t-s)</i>	- 0,080 **							
<i>Dakota du Nord(t-s)</i>	1,603 **							
<i>Tennessee(t-s)</i>	0,727 **							
<i>Utah(t-s)</i>	2,407 **							
<i>Wyoming(t-s)</i>	- 4,101 **							
$\Delta(\text{Tennessee})$	- 0,103 **							
$\Delta(\text{Utah})$	- 0,218 **							
$\Delta(\text{Wyoming})$	0,344 **							
$\Delta(\text{Dakota du Nord})$	- 0,144 **							
	R ² ajusté	0,931	R ² ajusté	0,963	R ² ajusté	0,973	R ² ajusté	0,996

* indique un coefficient statistiquement significatif au niveau de confiance 95%

** indique un coefficient statistiquement significatif au niveau de confiance 99%

* indique un coefficient statistiquement significatif au niveau de confiance 95%

** indique un coefficient statistiquement significatif au niveau de confiance 99%